Aus dem Institut für Veterinärbiochemie und Molekularbiologie
der Veterinärmedizinischen Fakultät
der Universität Zürich

Direktor: Prof. Dr. U. Hübscher

Co-expression of a functional clamp loader complex in *Escherichia coli*:

Construction of an expression operon for the five subunits of human replication factor C

INAUGURAL-DISSERTATION

zur Erlangung der Doktorwürde der
Veterinärmedizinischen Fakultät
der Universität Zürich

vorgelegt von

Sibylle Kuonen

Tierärztin von Zürich

genehmigt auf Antrag von
Prof. Dr. U. Hübscher, Referent
Prof. Dr. H. Lutz, Korreferent

Zürich 2003

Druck: Kantonale Drucksachen-
und Materialzentrale Zürich

Für meine Eltern
Helga und Markus
und meinen Bruder Sven

Content

page

1 Summary

1.1 English

In this thesis project a system for the overexpression of recombinant human replication factor C (RF-C) in an *E. coli* host was developed and optimised.

Cloning: A multi-step cloning strategy was designed to clone all five human RF-C genes into one vector. First, the cDNAs of the five RF-C genes were amplified by PCR and cloned into pET vectors. In several cloning steps three RF-C genes were cloned into a pET41a vector and the other two into a pET21a vector. All five genes were subsequently cloned into the pET21a vector under the control of the strong T7 promoter.

Expression: The pET21a/p38/p140/Hisp37/Hisp40/p36 vector construct was expressed in an *E. coli* BL21(DE3)-Codon Plus RIL strain. This *E. coli* strain additionally overexpresses tRNAs which recognise codons rarely used in *E. coli* (R = argU: AGA or AGG; I = ileY: ATA; L = leuW: CTA) to prevent codon-bias. The major obstacles during the expression of RF-C were the low expression level, the insolubility of the target protein and the rapid degradation of the large RF-C subunit (p140). The induction conditions had to be thoroughly optimised to obtain a reasonable amount of soluble target protein. The induction temperature had a very strong influence on the properties of recombinant RF-C, and lowering it to 23 °C dramatically increased the amount of soluble, full-length RF-C.

Purification: The binding ability of recombinant RF-C to two different materials was tested, to phosphocellulose and to nickel, because two of the RF-C subunits were 6 x His-tagged (Hisp40 and Hisp37). Strong binding of recombinant RF-C to phosphocellulose was observed. On the other hand, the binding affinity to nickel was only weak, but very specific. The two materials were therefore combined to a two step purification scheme, where the phosphocellulose purification step was immediately followed by the nickel purification step.

Enzyme activity: RF-C activity of the purest RF-C fraction was tested in DNA polymerase holoenzyme assays on a circular and a linear DNA template, respectively. Additionally, the ATP-dependent clamp loading activity of recombinant RF-C was confirmed in a [γ^{32}P]-PCNA loading assay onto circular double-stranded DNA. In all activity assays recombinant RF-C displayed the same enzymatic properties as it is known for RF-C purified from natural sources.

Finally, it was shown that bacterially overexpressed RF-C was not contaminated by bacterial exo- or endonucleases to assure that the quality of recombinant RF-C was sufficient for its future application in *in vitro* replication or repair activity assays which contain DNA.

1.2 Deutsch

Im Rahmen dieses Dissertationsprojektes wurde ein System für die Überexprimierung von rekombinantem menschlichen Replikationsfaktor C (RF-C) in *E. coli* entwickelt und optimiert.

Klonierung: Eine mehrstufige Klonierungsstrategie wurde entworfen, um alle fünf menschlichen RF-C Gene in einen Vektor zu klonieren. Zuerst wurden die cDNAs der fünf RF-C Gene mit PCR amplifiziert und in pET Vektoren kloniert. In mehreren Klonierungsschritten wurden drei RF-C Gene in einen pET41a Vektor, die anderen zwei in einen pET21a Vektor kloniert. Alle fünf Gene wurden anschliessend in den pET21a Vektor kloniert, wo sie unter der strikten Kontrolle des T7-Promoters standen.

Expression: Das pET21a/p38/p140/Hisp37/Hisp40/p36 Vektorkonstrukt wurde in einem *E. coli* BL21(DE3)-Codon Plus RIL Stamm exprimiert. Dieser *E. coli* Stamm überexprimiert zusätzlich tRNAs, welche von *E. coli* nur selten verwendete Codons erkennen (R = argU: AGA oder AGG; I = ileY: ATA; L = leuW: CTA), um dem Problem des Codon-bias (Bevorzugung gewisser Codons) vorzubeugen. Die grössten Hindernisse während der Expression von RF-C stellten der niedrige Grad der Expression, die Unlöslichkeit des Zielproteins und die schnelle Degradation der grossen RF-C Untereinheit (p140) dar. Die Induktionsbedingungen mussten sehr gründlich optimiert werden, um eine akzeptable Menge an löslichem Zielprotein zu erhalten. Die Induktionstemperatur hatte einen sehr starken Einfluss auf die Eigenschaften von rekombinantem RF-C, und ihre Senkung auf 23 °C erhöhte die Menge an löslichem und undegradiertem RF-C dramatisch.

Reinigung: Die Bindungseigenschaften von rekombinantem RF-C an zwei verschiedene Materialien wurden getestet, an Phosphozellulose und an Nickel, weil zwei der RF-C Untereinheiten ein 6 x His-tag trugen (Hisp40 und Hisp37). Eine starke Bindung von rekombinantem RF-C an Phosphozellulose wurde beobachtet. Im Gegensatz dazu war die Bindungsaffinität zu Nickel nur schwach, dafür war diese Bindung sehr spezifisch. Die beiden Materialien wurden darum zu einem Zwei-Stufen-Reinigungsschema kombiniert, indem auf den Phosphozellulose- sofort der Nickel-Reinigungsschritt folgte.

Enzymaktivität: Die RF-C-Aktivität der reinsten RF-C-Fraktion wurde mittels “DNA Polymerase holoenzyme Assays” auf einer zirkulären und einer linearen DNA-Vorlage getestet. Zusätzlich wurde die ATP-abhängige “Clamp-loading-Aktivität” von rekombinantem RF-C in einem “[γ^{32}P]-PCNA loading Assay” mit zirkulärer, doppelsträngiger DNA bestätigt. In allen Aktivitätsversuchen zeigte der rekombinante RF-C die gleichen enzymatischen Eigenschaften, wie sie von aus Säugerzellen gereinigtem RF-C bekannt sind.

Schliesslich wurde gezeigt, dass bakteriell überexprimierter RF-C nicht mit bakteriellen Exo- oder Endonukleasen kontaminiert war, um sicher zu gehen, dass die Qualität von rekombinantem RF-C für die zukünftige Anwendung in *in vitro* Replikations- und Reparaturversuchen, welche DNA enthalten, genügte.

2 Introduction

2.1 Genome integrity and cell division

The DNA of any living organism contains all the necessary information for the key characteristics of life. Successful function and reproduction are the main targets on a macroscopic level. On a molecular level a cell is faced with the crucial task of maintaining the stability of its genome and replicating its DNA. An enormous machinery of proteins coordinates these fundamental and complicated events. Despite the evolutionary distance the functions of proteins involved and the mechanisms at the basis of these events are strikingly similar in prokaryotes and eukaryotes. Even more complex is the situation in multicellular organisms. An intact genome is essential for the communication amongst cells. In order to form tissues they have to develop in an organised manner, divide at the right time and die if necessary. The loss of the ability of cells to communicate and act selflessly as a part of an integrated whole can lead to severe conditions, such as cancer and hereditary diseases.

Throughout evolution many mechanisms have been established to guarantee genome stability from generation to generation. Firstly, several repair mechanisms can bring damaged DNA back into its original state. They include base excision repair, nucleotide excision repair, mismatch repair and double-strand break repair. Secondly, DNA replication is performed by DNA polymerases and auxiliary proteins with high fidelity. Finally, the cell possesses very sophisticated cell cycle checkpoint pathways to survey the different phases of the cell cycle including cell division.

2.2 DNA replication: The proteins at the replication fork

2.2.1 DNA polymerases

DNA polymerases (pols) are the key enzymes to maintain genome stability. They are involved in many fundamental processes of a cell such as DNA replication, DNA repair, bypassing of damaged DNA and DNA recombination. All pols share a common activity, they catalyse the synthesis of DNA macromolecules from deoxyribonucleoside triphosphates using the chemical energy derived from the hydrolysis of the latter. At least 19 different pols are known in eukaryotes (reviewed in Hubscher, et al., 2002). Three of the best characterised pols have distinct functions at the replication fork (see Table 1). Pol α/primase associates with the initiation complex at the origin of replication and starts to synthesise a short RNA/DNA hybrid of approximately 10 RNA nucleotides followed by 20 – 30 DNA nucleotides. Then pol δ or pol ε takes over DNA synthesis and uses the RNA/DNA product of pol α/primase as a primer for processive DNA elongation. The event where pol α/primase is substituted by pol δ or ε is called polymerase switch and is coordinated by several auxiliary repli-

cation proteins including replication factor C (RF-C) (Maga, et al., 2000, Mossi, et al., 2000). Once DNA synthesis is initiated at the origin of replication it has to be continued along the DNA double helix. But this double helix is orientated in an antiparallel manner. Determined by the structure of the pol, which can catalyse DNA synthesis exclusively in the 5' → 3' direction, only one strand can be synthesised continuously and is called leading strand. The opposite mother strand (called lagging strand) needs to be synthesised in numerous short pieces called Okazaki fragments. The two DNA strands form a Y-shaped figure at the point of replication and this arrangement is hence called replication fork. It was found that pols at the replication fork could dimerise suggesting that the replication complex moved unidirectionally along the DNA (Podust, et al., 1992). To solve the sterical problem of antiparallel DNA synthesis on double-stranded DNA a model was proposed where the lagging strand formed a loop and allowed the pols on both leading and lagging strand to overall progress in the same direction (Waga and Stillman, 1994). The main task of the processive pols is to synthesise DNA with high speed and fidelity. They contain an additional 3' → 5' exonuclease activity to hydrolyse misincorporated nucleotides before they continue with DNA synthesis. This activity helps to minimise the mistake rate of processive pols and therefore increases their fidelity.

Pols are extremely specialised enzymes and consequently depend on the assistance of numerous auxiliary proteins to accomplish highly processive, accurate and coordinated DNA synthesis. They are the core enzymes of the eukaryotic replisome, but DNA replication has to be understood as the concerted action of many enzymes where manifold interactions between the proteins involved play an indispensable role in the correct and precise function of the whole complex (reviewed in Hubscher, et al., 2002).

Table 1 Functions of the classical eukaryotic DNA polymerases and telomerase [1]

Polymerase	No of subunits	Functional tasks
Pol α/primase	4	Initiator pol
Pol β	1	Base excision repair pol
Pol γ	2	Mitochondrial replication pol
Pol δ	4	Main pol at the leading and lagging strand
Pol ε	4	Leading and lagging strand
Telomerase	?	Telomere maintenance pol

[1] reviewed in Hubscher, et al., 2002, where also the novel pols are reviewed

2.2.2 Auxiliary proteins in DNA replication

Auxiliary proteins carry out many functions throughout all stages of DNA replication. Focussing on the replication fork the five main tasks are:

1. To relax the coiled DNA double helix ahead of and behind the replication fork. This step is catalysed by DNA topoisomerases I and II (reviewed in Wang, 2002).
2. To unwind the DNA double-strand. DNA helicases are involved in the formation of single-stranded DNA (reviewed in Caruthers and McKay, 2002, van Brabant, et al., 2000).
3. To stabilise single-stranded DNA: Single-stranded DNA is more prone to degradation than double-stranded DNA and has therefore to be stabilised by a single-strand binding protein called replication protein A (RP-A) (reviewed in Wold, 1997).
4. To coordinate the polymerase switch and tether the processive pols to the DNA: The ring-shaped trimeric protein proliferating cell nuclear antigen (PCNA) forms a clamp around the DNA and can be loaded onto and removed from the DNA by the clamp loader replication factor C (RF-C) (reviewed in Maga and Hubscher, 2003).
5. To process Okazaki fragments: Since lagging strand DNA synthesis is performed in small pieces, their ends have to be adapted and ligated. The two main enzymes involved in that step are flap endonuclease 1 (Fen 1) and DNA ligase I (reviewed in Henneke, et al., 2003a).

2.2.3 Proliferating cell nuclear antigen (PCNA) and replication factor C (RF-C)

PCNA is a ring-shaped trimeric protein, which has the ability of forming a clamp around the DNA. PCNA was found to interact with many proteins including pol δ and ε (Jonnson and Hubscher, 1997, Kelman, 1997, Krishna, et al., 1994, Ohta, et al., 2002, Schurtenberger, et al., 1997, Tsurimoto, 1999). The function of this interaction was soon apparent: The ring formed by PCNA tethers the processive pols to the DNA contributing to the speed and high processivity (Fukuda, et al., 1995, Wyman and Botchan, 1995). Several cell cycle checkpoint proteins also interact with PCNA and modulate the protein making it a perfect platform which is linked to the effectors of DNA replication, the polymerases, as well as to regulators of the cell cycle (Fotedar and Fotedar, 1995, Fujise, et al., 2000, Jonnson, et al., 1998, Maga, et al., 1997, Oku, et al., 1998, Sever-Chroneos, et al., 2001).

Since PCNA is a ring-like protein, it cannot move freely onto and from DNA and has to be loaded by its clamp loader, replication factor C (RF-C) (reviewed in Mossi and Hubscher, 1998). RF-C was first found to be a processivity factor for DNA synthesis by pol δ and ε selectively, but only in the presence of PCNA, RP-A and ATP (Fairman, et al., 1989, Tsurimoto and Stillman, 1989). Its interaction with PCNA was discovered, its function as the clamp loader of PCNA described and then further characterised (Podust, et al., 1995). RF-C is a pentameric complex consisting of one large subunit (p140) and four small subunits (p40,

p38, p37 and p36). The large subunit and several of the small subunits bind PCNA (see Table 2). RF-C also possesses an ATPase activity and a DNA binding activity, p37 specifically binds to primer-template junctions and p140 can also bind DNA. These different properties were characterised by numerous authors (Cai, et al., 1997, Cai, et al., 1998, Chen, et al., 1992a, Chen, et al., 1992b, Ellison and Stillman, 1998, Keller, et al., 1999, Kolpashchikov, et al., 2001, Mossi, et al., 1997, Neuwald, et al., 1999, Pan, et al., 1993, Podust, et al., 1998). Footprinting experiments revealed that at a template-primer junction RF-C covers 15 bases of the primer DNA from the 3'-end and 20 bases of the template (Tsurimoto and Stillman, 1990). The combination of all these activities led to the following model: RF-C can slide along the DNA until it encounters a primer-template junction, e.g. a RNA/DNA hybrid synthesised by pol α/primase or a nicked DNA template during DNA repair. In the absence of ATP RF-C binding to the junction is weak, addition of ATP increases the specific primer-binding activity (Tsurimoto and Stillman, 1991). Then PCNA binds to the RF-C-DNA complex, forming an unstable intermediate complex. PCNA loading is absolutely ATP-dependent, but ATP hydrolysis is not required. Upon ATP hydrolysis by the RF-C ATPase the protein-DNA complex undergoes a conformational change, which finally leads to the formation of the pol holoenzyme complex, consisting of pol δ and ε, PCNA and RF-C.

Table 2 Properties of the five human RF-C subunits [1]

	size (kDa)	PCNA binding	DNA binding	ATPase activity	Pol binding
p140	128.3	+	+		
p40	39	+		+	pol δ
p38	40.5	+			
p37	39.6		+ (at primer ends)	+	pol ε
p36	38.5	+		+	

[1] reviewed in Jeruzalmi, et al., 2002, Mossi and Hubscher, 1998

The function of a clamp sliding along the DNA which serves as a platform for coordinated DNA replication is conserved throughout evolution (reviewed in Jeruzalmi, et al., 2002). In *E. coli* the pol III holoenzyme, a ten-subunit complex, is responsible for DNA replication. It contains the DNA polymerase (three subunits), the clamp loader which is called γ complex (five subunits), the β clamp (a homodimer) and an additional connector protein called τ. This protein can bind two pol enzymes and one γ complex to form a tight asymmetric, dimeric pol complex (Jeruzalmi, et al., 2001a, Jeruzalmi, et al., 2001b, Kuriyan and O'Donnell, 1993, Leu and O'Donnell, 2001, O'Donnell, et al., 2001). The clamp loader complexes of both pro-

karyotes and eukaryotes require ATP for efficient clamp loading. Moreover, they show a high sequence similarity what suggests that their clamp loading mechanism may be very similar, too. Comparing the structure of the sliding clamps, PCNA in eukaryotes and the β clamp in *E. coli*, there is an obvious difference. The PCNA clamp contains three subunits, the β clamp only two. But the size of one subunit of PCNA is only about 70 % of the size of a subunit in the β clamp, so that three monomers of PCNA are necessary to form the typical ring-shaped clamp whereas only two β monomers are required for a functional clamp. The overall shape, however, is almost identical, even though the amino-acid identity is very low (< 5 %) suggesting a functional but not a primary-structural conservation over 10^9 years.

The sliding clamps and clamp loader complexes of many organisms including T4 bacteriophage, several bacteria and eukaryotes, e.g. yeast and humans, were identified and characterised (Bacteria: Kelman and Hurwitz, 2000, Pisani, et al., 2000. Yeast: Burgers, 1991, Cullmann, et al., 1995, Fien and Stillman, 1992, Gomes and Burgers, 2001, Gomes, et al., 2001, Hingorani and Coman, 2002, Holmes and Haber, 1999, Reynolds, et al., 1999, Schmidt, et al., 2001a, Schmidt, et al., 2001b, Yoder and Burgers, 1991. Human and bacteriophage T4: Tsurimoto and Stillman, 1990). They all show very high similarity in function, but slight differences in the subunit composition (see Table 3).

Table 3 Composition of the sliding clamp and its clamp loader in prokaryotes and eukaryotes [1]

		Prokaryotes		Eukaryotes	
		T4 bacteriophage	***E. coli***	***S. cerevisiae***	**mammal**
sliding clamp		gene 43 protein (homotrimer)	β (homodimer)	PCNA (homotrimer)	PCNA (homotrimer)
clamp loader	large subunit	gene 62 protein	δ	RF-C1 *(94.9 kDa)*	RF-Cp140 *(128 kDa)*
	small subunits	gene 44 proteins (homotetramer)	γ, δ', χ and ψ	RF-C4 *(36.2 kDa)* RF-C5 *(39.9 kDa)* RF-C2 *(39.7 kDa)* RF-C3 *(38.2 kDa)*	RF-Cp40 *(39 kDa)* RF-Cp38 *(40.5 kDa)* RF-Cp37 *(39.6 kDa)* RF-Cp36 *(38.5 kDa)*

[1] reviewed in Jeruzalmi, et al., 2002

The clamp loader structures from different organisms are compared in Table 3. It is always composed of a large and four small subunits. The homology of the individual subunits can be further visualised by amino-acid-sequence alignment. It revealed regions of striking similarity between prokaryotes and eukaryotes. The small RF-C subunits from human and yeast and the prokaryotic homologues align with the central part of the large RF-C subunit. These regions were termed RF-C boxes II – VIII (Allen, et al., 1998, Fotedar, et al., 1996). The large subunit contains an additional box (termed box I) of about 90 amino-acids within its N-terminal region. This region also shares regions of high similarity with DNA ligases and to a lesser extent with the eukaryotic poly (ADP-ribose) polymerases (PARP) and has been shown to bind DNA (Burbelo, et al., 1993, Fotedar, et al., 1996, Rheaume, et al., 1997). It was also identified as a distant version of the BRCT motif, which was first described as the C-terminus of the breast cancer protein BRCA1, a p53 binding protein. A superfamily of proteins shares this motif, the common property of which appears to be the participation in DNA-damage-responsive checkpoints (see chapters 2.3 and 2.4) (Bork, et al., 1997).

The boxes II – VIII are found in all five RF-C subunits. Their degree of conservation varies and they contain important catalytic sites as summarised in Figure 1 and Table 4. In all five subunits the complex formation domain is located at the C-terminal end of the protein where the amino-acid sequence is unique to each subunit (Cullmann, et al., 1995, Uhlmann, et al., 1997a, Uhlmann, et al., 1997b). On the contrary the N-terminal ends of the subunits contain the homology boxes and seem to be the “business-end” of each protein with multiple catalytic sites. Although the four small subunits are very similar to each other, they are not interchangeable and, despite their sequence redundancy, they are all required for the formation of an active RF-C complex.

Figure 1 Eukaryotic RF-C boxes and location of functional sites

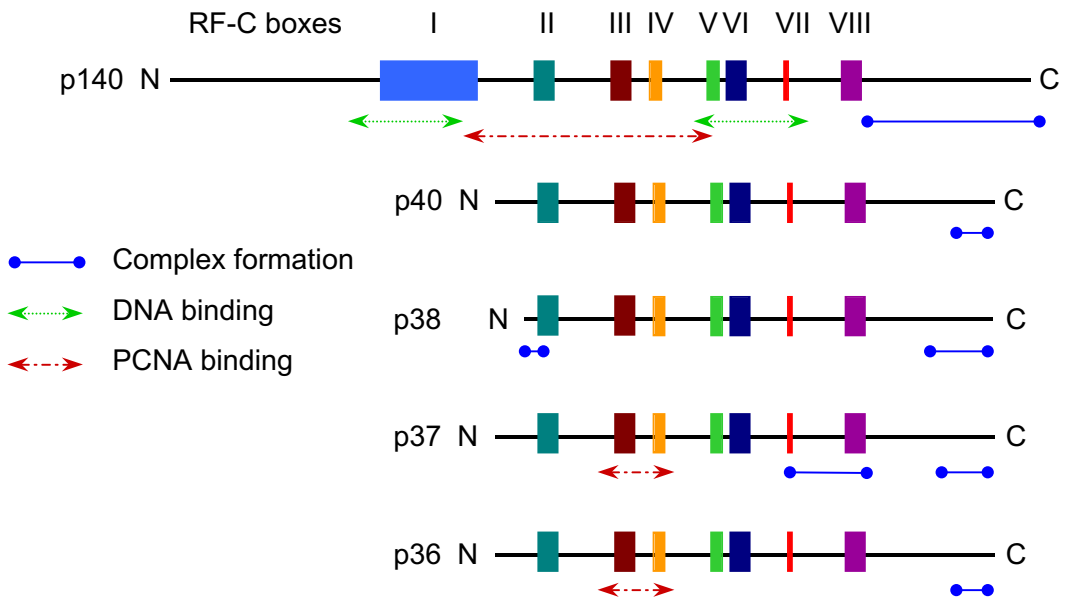


Table 4 Functions of the eukaryotic RF-C boxes [1]

	RF-C subunit	activity	similarity to
box I	p140	DNA binding	DNA ligases, poly (ADP-ribose) polymerase
box II	all	ATP/GTP binding	unique
box III	all	phosphate binding (P-loop)	
box IV	all	ATP/GTP binding	unique
box V	all	?	DEAD-box proteins (putative RNA helicases with P-loops and ATPase activity)
box VI	all	ATP/GTP binding	unique
box VII	all	?	?
box VIII	all	?	?

[1] reviewed in Jeruzalmi, et al., 2002, Mossi and Hubscher, 1998

2.3 Cell cycle regulation

Replication of the genomic DNA is the key event of cell division, but a large number of regulatory factors are required to guide the cell through a coordinated cell cycle. The cell cycle is divided into four phases: G1-phase (gap 1), the phase of the actual cell activity. This phase is followed by the S-phase (synthesis of DNA) where the genomic DNA is copied. The third phase is called G2-phase (gap 2), a short interval in which cell division is prepared and initiated. The M-phase (mitosis) unites all the different stages of mitosis: Condensation of the chromosomes, alignment of the homologous chromosomes in the central plane of the cell, separation of the chromosomes into two equivalent sets of genomic DNA copies and finally cell division. Differentiated cells which cannot divide any longer proceed into an additional phase called G0 (gap 0).

As the cell proceeds through the cell cycle the transitions from one phase to the following phase are regulated by cyclin-dependent kinases (reviewed in Nigg, 1995). Cyclins are periodically expressed proteins which can bind to a constantly expressed kinase and activate its kinase activity. The complex is subsequently able to phosphorylate target proteins. In case of the onset of DNA replication or in other words the G1/S-phase transition two cyclin-dependent kinases, cyclinE/Cdk2 and cyclinA/Cdk2, play an important role. The activity peak of cyclinE/Cdk2 is found in late G1-phase, before DNA synthesis is initiated, while cyclinA/Cdk2 activity appears later, at the onset of DNA replication (reviewed in Fotedar and Fotedar, 1995). In eukaryotes this initiation step can be divided into two main events. First, in late G1-phase a preinitiation complex containing ORC proteins is formed at the origin of replication.

Then, during the G1/S-phase transition, cyclin-dependent kinases promote the conversion of the preinitiation complex into an active replication form through recruitment of numerous proteins, including RP-A, pol α/primase, RF-C, Fen 1 and DNA ligase I. Cyclin-dependent kinases also appear to have a negative effect on DNA replication by catalysing a series of phosphorylation events which lead to the inability of initiation factors to re-enter the preinitiation state, hence preventing re-replication of the genome (reviewed in Henneke, et al., 2003b).

Recently it has been shown that cyclin-dependent kinases are associated with the human DNA synthesome, a replication competent multiprotein-complex containing pol α/primase, pol δ, RP-A, DNA ligase I, DNA topoisomerase I and II, DNA helicase, PCNA, RF-C and poly (ADP-ribose) polymerase. Following the DNA synthesome through the cell cycle several different combinations of cyclin-dependent kinases were present in this multiprotein-complex, suggesting a regulatory influence on the function of the synthesome, but the precise mechanisms remain to be elucidated (Frouin, et al., 2002).

2.4 DNA damage checkpoint and repair

RF-C plays an important role and has versatile functions in the DNA damage checkpoint and repair pathways. The DNA damage checkpoint is a signal transduction pathway which can block cell cycle progression or slow down the rate at which S-phase is passed through. It is a surveillance mechanism and is usually not required for cell cycle events, but it enforces their proper order, which is especially critical after acute DNA damage or replication error. DNA damage checkpoint proteins are generally well conserved amongst eukaryotes apart from some exceptions in the checkpoint signalling pathways (reviewed in Canman, 2001, Melo and Toczyski, 2002). The signalling cascade starts with the recognition of DNA lesions. Two multiprotein complexes recognise damaged DNA independently and can activate the checkpoint pathway. In mammalian cells these are firstly a pair of large protein kinases with homology to phosphatidylinositol-3-kinases (PI3K-like checkpoint kinases: ATM = Ataxia telangiectasia mutated and ATR = Ataxia and Rad-related) which are able to associate with damaged DNA and initiate a phosphorylation cascade. Secondly three checkpoint proteins (Rad9, Rad1 and Hus1, called the 9-1-1-complex) with sequence similarity to PCNA and Rad17, another checkpoint protein with sequence similarity to the large subunit of RF-C (p140), were discovered (Lindsey-Boltz, et al., 2001, Rauen, et al., 2000, Roos-Mattjus, et al., 2002, Shiomi, et al., 2002). Mass spectrometric analysis of the Rad17-homologue complexes in yeast *S. cerevisiae* showed that the Rad17-homologue replaces the large RF-C subunit and exists in a complex with the other four RF-C subunits. Moreover, Rad17 binds the 9-1-1-complex directly. These studies led to the suggestion that the 9-1-1-complex might be loaded onto damaged DNA by the Rad17-RF-C-complex in analogy to the interactions of

PCNA and RF-C (Venclovas, et al., 2002). The structural similarities of the 9-1-1-complex to the PCNA clamp and of the Rad17-RF-C-complex to the classical RF-C-clamp-loader were confirmed by glycerol spray / low voltage microscopy (Griffith, et al., 2002).

It is not yet known whether the two types of damage recognition complexes associate with naked DNA or with DNA-protein complexes, e.g. a stalled replication machinery. However, they could subsequently recruit adapter proteins and effector kinases to the site of damage (reviewed in Canman, 2001, Melo and Toczyski, 2002). The 9-1-1-complex might serve as a landing pad for the recruited proteins assembling them at the site of the DNA lesion and localising them close to the sensor kinases as substrates for phosphorylation. These downstream pathways are so far only poorly characterised in mammalian cells, but two checkpoint kinases (Chk1 and Chk2) were found to be effector kinases which are phosphorylated and activated upon DNA checkpoint pathway activation. This activation step was found to require one of two even less characterised adapter proteins. The adapters appear to function by promoting checkpoint signalling from the sensor kinases to the effector kinases in a manner dependent on the type of DNA damage present. The two proposed human adapter proteins are claspin and the breast cancer protein BRCA1.

The DNA damage checkpoint pathway targets many fundamental responses of a cell to damage of its genomic DNA, not only, as mentioned above, blockage or delay of the cell cycle, but also induction of the transcription of repair genes, promotion of repair processes, stabilisation of stalled replication forks and – when the integrity of the genome cannot be ensured – apoptosis.

When a DNA lesion is recognised and marked in a cell, it can be repaired by several mechanisms depending on the type of damage. Although different kinds of lesions require specific mechanisms of repair and exquisite specificity is found for certain proteins in the recognition of DNA damage, there is often little or no specificity for the enzymes required for the resynthesis step. On the other hand the repair pathways show redundancy and can even replace each other when one pathway is partially defective. Damaged bases can be replaced either by base excision repair (BER) or by nucleotide excision repair (NER). In mammalian cells BER is initiated by DNA glycosylases, which remove defective bases and therefore create an apurinic site. Endonucleases then cut out this single apurinic nucleotide and pol β fills in the correct nucleotide. An alternative pathway is known to process apurinic sites: A small patch of about 7 nucleotides is excised and the gap filled in a PCNA- and RF-C-dependent manner by pol δ or ε (Pascucci, et al., 2002, reviewed in Matsumoto, 2001).

On the contrary, NER in mammalian cells completely depends on PCNA and RF-C, suggesting the participation of pol δ or ε in this pathway where the DNA lesion is spatially excised. Both pols are able to perform repair DNA synthesis. They can efficiently substitute for each other in case one is unavailable (Shivji, et al., 1995, reviewed in Wood and Shivji, 1997).

DNA lesions not only include defective but also misincorporated nucleotides, which directly lead to mutagenesis. It is therefore a crucial task of the replication machinery to reduce the error rate during replication to an absolute minimum. The error rate of a processive DNA pol (e.g. pol δ or ε) is estimated at 1 misincorporated nucleotide per 10^5 nucleotides incorporated. This value includes the intrinsic proof-reading activity (3' $\rightarrow$ 5' exonuclease). Considering that the human genome counts 10^9 base pairs a total of 10^4 nucleotides would be misincorporated in each cell cycle. DNA polymerase accessory proteins like PCNA and RF-C are known to increase the fidelity and processivity of the pol holoenzyme compared to the core polymerase up to 1000 fold. Finally, the cell possesses a system of postreplicative repair enzymes, which recognise nucleotide mismatches and correct the newly synthesised DNA strand (Kolodner and Marsischky, 1999). The exact mechanisms of these enzymatic reactions are not yet known in eukaryotic cells. All these mechanisms contribute to an extremely low error rate during DNA replication of $1 : 10^{10}$. In other words this figure expresses the fact that statistically a spontaneous mutation occurs in every tenth cell cycle.

2.5 RF-C p140 as a possible transcription factor

Recently, RF-C p140 has been reported to interact with transcription factors and tumour suppressor proteins, such as the retinoblastoma tumour suppressor protein, the tumour suppressor p53, the transcription factor C/EBPα, the histone deacetylase 1 (HDAC1), a BRCA1 complex and the Rel A subunit (p65) of the transcription factor NF-κB (Anderson and Perkins, 2002, Anderson and Perkins, 2003, Hong, et al., 2001, Pennaneach, et al., 2001). For Rel A it was also shown that p140 stimulates its transactivation activity *in vivo*, but it remained unclear whether this effect was caused by p140 alone or in a complex with the four small RF-C subunits. RF-C p140 was subsequently proposed to be a component of a transcriptional co-activator complex which selectively regulates the expression of genes controlled by NF-κB and other transcription factors. Since NF-κB acts as an important regulator of inflammation, proliferation and apoptosis the significance of this interaction awaits further investigation.

3 Problem and approach

3.1 History of recombinant RF-C

RF-C was first discovered and isolated from a human cell line (293 cells) in 1989 (Fairman, et al., 1989, Tsurimoto and Stillman, 1989). It was termed replication factor C after its function of an essential replication factor for *in vitro* simian virus 40 DNA replication and its property to elute from phosphocellulose in fraction C using the original purification protocol. Later RF-C was also purified from HeLa cells, calf thymus and yeast and it soon became obvious that RF-C was a very difficult protein-complex to purify, with loss of activity, degradation of the five subunit complex and low yield of pure protein being the main obstacles. As RF-C was found to be involved in manifold pathways of replication, cell cycle regulation, repair and DNA damage checkpoint, there was a growing demand for large amounts of highly purified, active RF-C to study its biochemical properties and possible function in all these pathways. The cDNAs of the five RF-C subunits were cloned and many different attempts have been made to produce active recombinant RF-C confirming the fact that RF-C was an extremely tricky and delicate enzyme, which had to be handled with special care. The different systems which had been tested for human and yeast RF-C were:

1. ***In vitro* coupled transcription / translation system:** The genes of all five human RF-C subunits, each cloned into a separate pET vector, were simultaneously transcribed and translated in a single *in vitro* reaction. The amount of synthesised product was low and the large subunit was partly degraded, possibly due to the fact that internal methionin codons were used as alternative start codons and therefore led to N-terminal degradation (Uhlmann, et al., 1996).

2. **Overexpression of yeast RF-C in yeast:** A yeast-vector containing the five genes of yeast RF-C was used to transfect a protease deficient yeast strain. After induction RF-C was overexpressed about 80 fold. Homogeneous and active yeast RF-C was subsequently purified (Burgers, 1999, Gerik, et al., 1997).

3. **Baculovirus overexpression system:** “High five” insect cells were infected simultaneously with five recombinant viruses which each produced one of the five human RF-C subunits. The difficulty in this system turned out to be the fact that statistically the insect cells are not infected with five viruses at an equimolar rate. Some cells might therefore lack one RF-C gene, but others might contain an extra copy of the same gene, leading to the production of incomplete and faulty RF-C complexes. Furthermore this system is an expensive and time-consuming method of producing recombinant proteins (Cai, et al., 1996, Podust and Fanning, 1997).

4. **Overexpression of individual subunits in *E. coli*:** All five human RF-C subunits were individually cloned into bacterial expression vectors (e.g. pET vectors) and overexpressed in an *E. coli* expression system. Unfortunately, all subunits turned out to be completely insoluble and formed so-called inclusion bodies. All attempts to solubilise the subunits and reconstitute the five subunit RF-C complex *in vitro* failed, and the insoluble RF-C subunits could only be used for antibody production (Chen, et al., 1992a, Chen, et al., 1992b, Luckow, et al., 1994).

5. **Overexpression of a truncated form of yeast RF-C in *E. coli*:** The five subunits of yeast RF-C were cloned into one bacterial expression vector (pBL 481). Overexpression was carried out through a BL21(DE3) *E. coli* strain in the presence of an additional plasmid (pSBETa) containing the information for the argU tRNA. In initial attempts to express full-length RF-C1 no polypeptide corresponding to RF-C1 was detectable after induction. Therefore a truncated form lacking 273 amino-acids from the N-terminus was chosen for overexpression of the five subunit complex. Simultaneous overexpression of the rare arginine tRNA (argU tRNA) as well as optimisation of the induction conditions drastically increased the yield of expressed RF-C (Gomes, et al., 2000).

In the need of large quantities of active human RF-C the goal of this thesis project was to develop and optimise a system for the overproduction and purification of active, full-length human RF-C, also taking advantage of recently developed bacterial expression systems which were improved for the production of difficult target proteins.

3.2 Translation issue problems

The frequency of codon-usage varies between organisms, and therefore the expression level of the corresponding tRNA also shows considerable differences. The problematic rare codons are known in *E. coli*. They include the arginine codon (argU: AGA or AGG), the isoleucine codon (ileY: ATA), the leucine codon (leuW: CTA) and finally the proline codon (proL: CCC). Efficient production of heterologous proteins in *E. coli* can be limited by the rarity of the tRNAs mentioned above, which are abundant in the organism from which the protein originated. The consequences for expression are transcription problems which lead to low or even lack of protein synthesis, early termination, misincorporation of amino-acids in the target protein and formation of insoluble protein aggregates (inclusion bodies). This codon-bias can be the main obstacle in the expression of difficult target proteins. In order to circumvent this problem new *E. coli* strains were designed which express rare tRNAs from an additional plasmid. They derive from an *E. coli* BL21 strain which lacks the Lon protease and the OmpT protease and are called *E. coli* BL21-Codon Plus RIL (extra copies of the argU,

ileY and leuW tRNA genes) and *E. coli* BL21-Codon Plus RP (extra copies of the argU and the proL tRNA genes). To make optimal use of these new *E. coli* strains the codon-usage of the proteins to be expressed has to be compared with the codon-usage of the host organism, in this case *E. coli*. The result of the analysis of the five human RF-C subunits and the RF-C complex in comparison with *E. coli* are shown in Table 5.

Table 5 Codon frequencies of selected codons in *E. coli* and the human RF-C genes.

gene(s) analysed		*E. coli* genome	p140	Hisp40	p38	Hisp37	p36	5 subunit human RF-C construct
no of codons		-	1149	375	357	384	341	2606
	AGG	1.4	8.7	**26.7**	**14.0**	0.0	8.7	11.5
Arg	AGA	2.1	**14.8**	2.7	**19.6**	**13.0**	**14.8**	**13.8**
	CGA	3.1	6.1	5.6	2.8	**18.2**	6.1	8.1
Leu	CTA	3.2	3.5	5.3	8.4	7.8	11.7	6.1
Ile	ATA	4.1	**14.8**	2.7	5.6	**13.0**	**14.7**	11.5
Pro	CCC	4.3	11.3	**16.0**	5.6	10.4	**17.6**	11.9

Codon frequencies are expressed as codons used per 10^3 codons encountered. Codons which might be critical for efficient expression of RF-C are printed in bold letters (value > 12 per 10^3 codons)

The most critical codon in all five RF-C subunits clearly is the argU codon (AGA or AGG). The leuW (CTA) codon is not critical in any of the RF-C subunits. For the ileY (ATA) and the proL (CCC) codon the situation is more complex. The former might be problematic for the expression of the RF-C p140, Hisp37 and p36 genes, whereas the latter might cause codon-bias during the translation of the RF-C Hisp40 and p36 gene product. However, this analysis of the RF-C genes indicates that the recently developed *E. coli* BL21-Codon Plus strains, either BL21-Codon Plus RIL (extra copies of the argU, ileY and leuW tRNA genes) or BL21-Codon Plus RP (extra copies of the argU and the proL tRNA genes), might promote high level expression of recombinant RF-C.

4 Materials and Methods

4.1 Plasmids

The plasmids containing the cDNAs of the five human RF-C subunits, pET16b/p140, YEp-Gal/Hisp40, pET19b/p38, pBlueskriptSK(+)/Hisp37 and YEpGalura/p36 were from K. Dahm (this Institute). The cloning / expression vectors, pET21a and pET41a, were purchased from Novagen.

4.2 Bacterial strains

An *E. coli* DH5α (Invitrogen) strain was used for plasmid amplification and during the cloning procedure. During the last step of cloning an *E. coli* XL10-Gold strain (Stratagene) was also used. For protein expression several different *E. coli* BL21(DE3) strains (Stratagene) were used, including BL21(DE3), BL21(DE3)Gold, BL21(DE3)-Codon Plus RIL, BL21(DE3)-Codon Plus RP and BL21(DE3)pLysS.

4.3 Agar plates and liquid medium for bacterial cultures

Liquid medium was prepared with LB-capsules (Q-BIOgene) according to the manufacturer's instructions, containing 10 g Tryptone-B, 5 g yeast extract-B and 10 g NaCl per 1 L of medium, and was then autoclaved. For the preparation of agar plates 15 g of Bacto™ agar (Becton, Dickinson and Company) was added to 1 L of liquid medium prior to sterilisation. Afterwards, the medium was cooled to 55 °C, the antibiotics or combinations of antibiotics added and the medium poured into Petri dishes. The following antibiotics were used during cloning and expression: For ampicillin (Amp, Fluka) a 70 mg/ml stock solution in water was prepared and was diluted to a final concentration of 70 μg/ml in agar plates and liquid medium. For kanamycin (Kan, Fluka) a 30 mg/ml stock solution in water was used and diluted to 30 μg/ml in the LB culture. Finally, a chloramphenicol (Cam, Sigma) stock solution of 34 mg/ml was prepared in ethanol and diluted to a concentration of 34 μg/ml in the medium. All antibiotic stock solutions were filter-sterilised prior to use. LB plates containing all possible combinations of these three antibiotics were prepared.

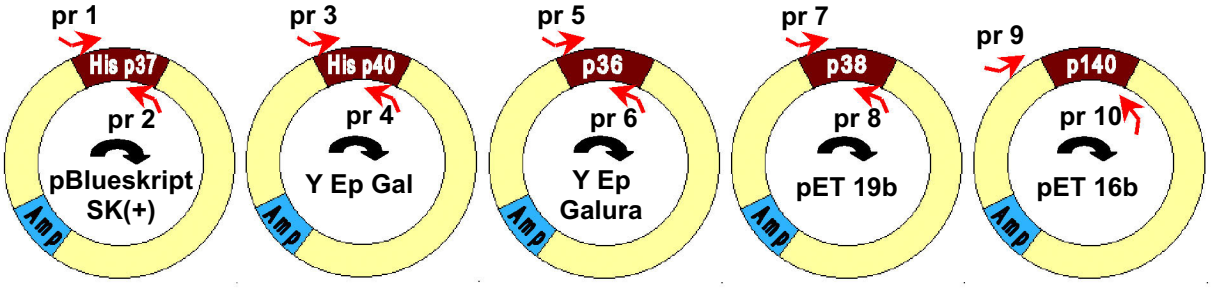
4.4 PCR amplification of the five human RF-C genes

For the amplification of each gene two primers were designed, which also contained the sites for restriction enzyme digestion at the 5'-end. In the case of the p140 gene, which was already cloned in a pET16b vector, primer sequences complementary to flanking parts of the gene on the vector were chosen, so that the p140 gene including the bacterial ribosome binding site upstream of the gene was amplified. In the downstream primer a sequence complementary to the bacterial ribosome binding site was added to the 5' primer end. Upon completion of the PCR reaction the p140 PCR product therefore carried two ribosomal binding sites, one on either side of the gene. The primers for the PCR amplification of the other four RF-C genes were designed to anneal to the gene sequences at the beginning and at the end of each gene, as shown in Table 6 and Figure 2 (all primers were ordered from Microsynth). Additionally, the commercial vector primers T7-promoter-primer and T7-terminator-primer (both from Novagen) were used for the PCR analyses of cloning products.

Table 6 Primer sequences used in this study

	Template	Sequence (5' → 3'), restriction sites and start / stop codons	No of bases
pr 1	His p37	AAA AAA ATT AAT CGG CCG **ATG** GGC AGC AGC CAT CA Ase I Eag I *Met*	35
pr 2	His p37	AAA AAA ACT AGT AAT TCC **TTA** ACA ATT CTG AGA Spe I *Stop*	33
pr 3	His p40	AAA AAA ATT AAT **ATG** GGC AGC AGC CAT CA Ase I *Met*	29
pr 4	His p40	AAA AAA ACT AGT **CTA** TGA GGC CAC CGG GGC CAT T Spe I *Stop*	34
pr 5	p36	AAA AAA CAT **ATG** GAG ACC TCA GCA CTC Nde I/*Met*	27
pr 6	p36	AAA AAA GGA TCC TCG AG**C TA**G GCC TCT GCA ACA AT BamH I Xho I *Stop*	35
pr 7	p38	AAA AAA CAT **ATG** ATG AGC CTC TGG GTG GA Nde I/*Met*	29
pr 8	p38	AAA AAA AAG TCG ACA G**TC A**GA ACA TCA TGC CTT Sal I *Stop*	33
pr 9	p140 (vector)	GGG GGG GGG TCG ACT TGT TTA ACT TTA AGA AGG AGA TAT Sal I	39
pr 10	p140 (vector)	GGG GGGCGG CCG ATT ATG TAT ATC TCC TTC TTT GTT AGC AGC... ...CGG GAT C Eag I	51

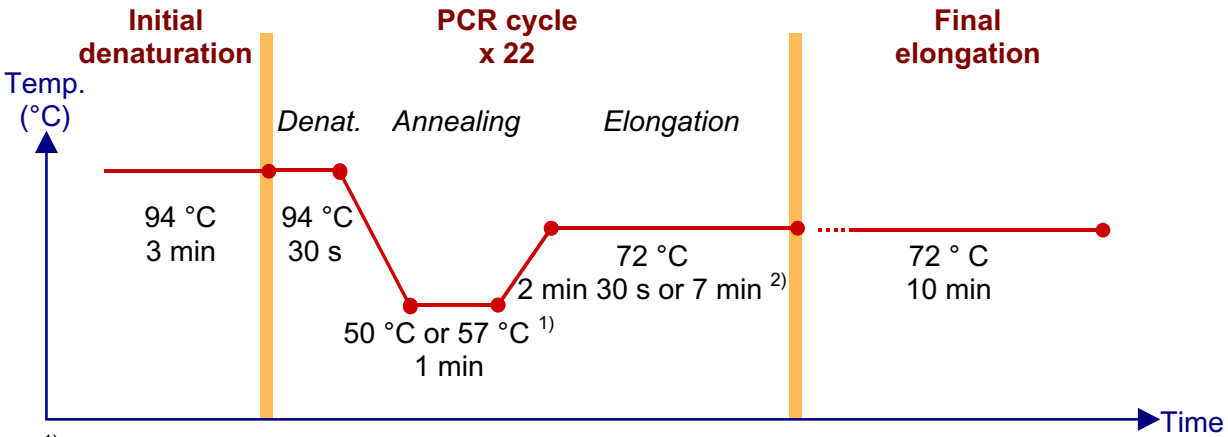
Figure 2 Primer design for the PCR amplification of the five human RF-C cDNAs



The PCR amplification mix contained 25 ng plasmid template, 25 pmol of each primer and 3 u of Pfu-polymerase (Promega). The PCR reaction was performed in a reaction volume of 50 µl using the reaction buffer provided by the manufacturer containing 20 mM Tris-HCl (pH 8.8), 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_2$, 0.1 % Triton®X-100 and 0.1 mg/ml nuclease-free BSA. Finally, dNTPs were added to a final concentration of 0.5 mM of each deoxyribonucleoside triphosphate. Two different PCR cycles were programmed, one for the gene of the large RF-C subunit and one for the genes of the small subunits (Figure 3 and Table 7).

- The annealing temperature was calculated according to the complementary sequence of each primer: **Temperature = (A + T) x 2 °C + (G + C) x 3 °C**
  5 °C were deducted for the annealing temperature of the PCR cycle.
- The elogation time was calculated according to the synthesis rate of the Pfu-polymerase, which is **1000 bp in 2 min**.

Figure 3 PCR cycle programs



1) see Table 7 under annealing temperature

2) see Table 7 under elongation time

Table 7 Conditions for PCR amplification

Gene	Primer	Primer length	Of which complementary	Annealing Temperautre calculated PCR program		Size of PCR product	Elongation time
Hisp37	pr 1 pr 2	35 bp 33 bp	17 bp 21 bp	54 °C 54 °C	**50° C**	1188 bp	**2 min 30 s**
Hisp40	pr 3 pr 4	29 bp 34 bp	17 bp 16 bp	54 °C 56 °C	**50 °C**	1143 bp	**2 min 30 s**
p36	pr 5 pr 6	27 bp 35 bp	18 bp 18 bp	56 °C 54 °C	**50 °C**	1052 bp	**2 min 30 s**
p38	pr 7 pr 8	29 bp 33 bp	17 bp 19 bp	54 °C 54 °C	**50 °C**	1099 bp	**2 min 30 s**
p140	pr 9 pr 10	39 bp 51 bp	25 bp 22 bp	62 °C 62 °C	**57 °C**	3542 bp	**7 min**

4.5 Cloning

4.5.1 Plasmid amplification and DNA preparation

A 5 ml LB culture containing the appropriate antibiotic(s) was inoculated with a single colony from an agar plate culture. The culture was incubated at 37 °C overnight. It was either directly used for DNA minipreparation or transferred into a 100 ml or 500 ml culture containing fresh LB medium and antibiotics. The culture was again incubated at 37 °C overnight. For midipreparation (100 ml culture) and maxipreparation (500 ml culture) the plasmid DNA was purified using a DNA preparation kit (Qiagen). For DNA minipreparations (5 ml culture) either a spin-column kit (Macherey-Nagel) was used if the purified plasmid was to be sequenced or the DNA was purified according to the following protocol. Buffers used are the same as for the midi- and maxipreparation protocol (Qiagen). The cells from 5 ml overnight culture were harvested by centrifugation (2700 g for 5 min). The bacterial pellet was resuspended completely in 250 µl of buffer P1. Then 250 µl of buffer P2 were added, the tubes inverted gently and incubated at room temperature for 5 min. 300 µl of cold buffer P3 were added, the tubes again inverted several times and incubated on ice for 10 min. The samples were centrifuged at 4 °C and 20,000 g for 10 min and the supernatant was carefully transferred into new tubes. 560 µl (0.7 volumes) of isopropanol were added to the supernatant, mixed carefully and incubated at room temperature for 5 min to precipitate the plasmid DNA. The DNA was then separated by centrifugation (4 °C, 20,000 g, 10 min), the DNA pellet washed with 800 µl of cold ethanol for 5 min (incubation on ice). Finally, it was centrifuged again for 5 min (4 °C, 20,000 g), the supernatant was discarded and the pellet left to dry. The plasmid DNA was dissolved in sterile, distilled water and stored at –20 °C until further use.

4.5.2 Preparation of competent cells and bacterial glycerol stocks

Competent cells: A 1 ml LB culture containing either the appropriate or no antibiotic was inoculated with the required bacterial strain and incubated at 37 °C overnight. The culture was then transferred into a 100 ml culture containing fresh LB medium and antibiotic. It was again incubated at 37 °C and OD_{600} was measured. When the culture reached an OD_{600}-value between 0.2 and 0.3, the cells were harvested by centrifugation at 4 °C and 3000 g for 5 min. The pellet was resuspended in 50 ml of cold, sterile $CaCl_2$ (50 mM) and left on ice for 20 min. The cells were centrifuged again in the same way as before and the pellet resuspended in 4 ml cold, sterile $CaCl_2$ (50 mM) and 750 µl cold, sterile glycerol (87 % (v/v)). The competent cells were divided into 200 µl aliquots, frozen in liquid nitrogen and stored at –80 °C until further use.

Glycerol stocks: A small LB-culture containing the appropriate antibiotic(s) was inoculated with one colony from an overnight plate culture and incubated at 37 °C overnight. 850 µl of the culture were mixed with 150 µl of cold sterile glycerol, the solution was immediately put on ice, then frozen in liquid nitrogen and stored at –80 °C until further use.

4.5.3 Transformation

100 µl of competent *E. coli* cells were mixed with 100 – 400 ng of plasmid or 10 µl of ligation mix and incubated on ice for 10 min. A 1 min heat shock at 42 °C was applied to the bacteria and then they were left on ice for another 2 min. 800 µl of LB-medium without antibiotic were added to the mix and it was subsequently incubated at 37 °C for 30 min. Finally, 100 µl of the mix were plated out on LB-plates containing the appropriate antibiotic(s). If a ligation mix was used for transformation, the mix was centrifuged at 100 g for 5 min, the bacterial pellet resuspended in 200 µl of LB-medium and half or all of this mix was plated out. The plates were incubated at 37 °C overnight.

4.5.4 Restriction enzyme digestion

During the cloning procedure the following restriction enzymes (all were purchased from New England BioLabs) were used, target site and enzyme activity of the enzyme stock are indicated in brackets: Nde I (5’ CA’TATG 3’, 20,000 u/ml), Ase I (5’ AT’TAAT 3’, 10,000 u/ml), BamH I (5’ G’GATCC 3’, 10,000 u/ml), Spe I (5’ A’CTAGT 3’, 10,000 u/ml), Xba I (5’ T’CTAGA 3’, 20,000 u/ml), Eag I (5’ C’GGCCG 3’, 10,000 u/ml), Sal I (5’ G’TCGAC 3’, 10,000 u/ml) and Xho I (5’ C’TCGAG 3’, 20,000 u/ml).

In a standard digestion reaction 1 µg of DNA was incubated for 2.5 hrs with one or two restriction enzymes (10 u of each enzyme) in 20 µl of the appropriate reaction buffer supplied by NEB and – if necessary – BSA was added to a final concentration of 100 µg/ml.

When PCR products where digested, the incubation time was increased to 15 hrs. For large plasmids the amount of DNA digested and the reaction volume was increased, but the enzyme concentration remained unchanged. These modified reactions are further described in “Results”.

The following reaction buffers supplied by the manufacturer were used:

NEB2: 10 mM Tris-HCl (pH 7.9), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT

NEB3: 50 mM Tris-HCl (pH 7.9), 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT

NEB4: 20 mM Tris-acetate (pH 7.9), 50 mM K-acetate, 10 mM Mg-acetate, 1mM DTT

NEBSal: 10 mM Tris-HCl (pH 7.9), 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT

4.5.5 Gel extraction

Digested DNAs were purified in two ways. First, a commercial DNA extraction kit (Qiagen) was used when the restriction reaction mix was loaded onto a 1 % agarose gel (in TAE buffer containing 0.5 μg/ml ethidiumbromide) and a specific DNA fragment had to be purified from a gel slice. Second, when in a two step digestion the reaction buffer of the first reaction had to be removed, the DNA extraction protocol was modified leaving out the agarose gel and treating the reaction mix in the same manner as a gel slice. On the contrary to the manufacturer’s protocol the DNA was eluted with 30 μl sterile, distilled water.

4.5.6 Ligation

The concentrations of the two or three restricted DNA fragments to be ligated were estimated according to the signal of the fragments on an agarose gel. The DNA fragments were mixed in a 1.5 : 1 insert : vector-ratio. For difficult ligation reactions several ratios between 1 – 2 : 1 were tested. The DNA was incubated with 400 cohesive end ligation units of T4 DNA ligase (New England BioLabs) in 20 μl of the ligation buffer (50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 μg/ml BSA) supplied by the manufacturer at 16 °C over-night.

4.6 Expression

4.6.1 Small scale induction and solubility test

The BL21(DE3) expression strains mentioned above were transformed either with the pET21a/p38/p140/Hisp37/Hisp40/p36 plasmid construct or with both the pET21a/p38/p140 and the pET41a/Hisp37/Hisp40/p36 constructs. A plate culture was incubated at 37 °C over-night, two colonies for each expression strain were picked and grown at 37 °C overnight in a small LB culture containing the appropriate antibiotics. A small sample of each culture was prepared as bacterial glycerol stock and kept for further analyses. The rest of the culture was

centrifuged at 800 g for 3 min, the bacterial pellet washed with fresh LB medium to remove even traces of β-lactamase from the medium, centrifuged again (800 g, 3 min) and finally resuspended in the same volume of fresh LB medium containing antibiotics. The culture was diluted 1 : 50 into a 50 ml LB culture containing the appropriate antibiotics. The culture was incubated at 37 °C and OD_{600} of the culture was monitored. Target gene expression was induced with IPTG when OD_{600} had reached a value of 0.6 or 1 and the temperature adjusted to either 23 °C, 30 °C or 37 °C. The concentrations of IPTG in the culture used for induction varied between 0.5 mM and 1 mM. In a standard small scale induction test the cells were induced at OD_{600} = 1 with 1 mM IPTG and incubated at 30 °C after induction. 2 ml culture samples were taken just before induction and then for 4 hrs every full hour after induction. OD_{600} of all samples was measured to monitor cell growth and target gene toxicity. Then, the cells of these samples were harvested by centrifugation (2700 g, 5 min, 4 °C) and the bacterial pellets were frozen at –20 °C for further analyses. After four hours the remaining 40 ml of culture were also centrifuged, the pellet frozen at –20 °C and kept for preliminary purification studies.

For the expression time curve and solubility analysis the cell pellet was resuspended in 300 µl lysis buffer A (see under "Purification"). The bacterial cells were cracked by sonication (level 4, 40 % duty cycle, 6 times 6 pulses) and, if the soluble faction had to be analysed, centrifuged at 20,000 g and 4 °C for 15 min. The supernatant was transferred into another tube and the pellet resuspended in the same volume of lysis buffer A. The protein concentrations of all the fractions (total cell protein, soluble and insoluble fractions) were calculated by Bradford assay. Then the bacterial proteins in each fraction were separated by SDS-PAGE and analysed either by Coomassie-blue staining or by Western blot analysis.

4.6.2 Medium and large scale induction

A 5 ml LB culture containing ampicillin and chloramphenicol was inoculated with the BL21(DE3)-Codon Plus RIL strain carrying the pET21a/p38/p140/Hisp37/Hisp40/p36 plasmid construct from a glycerol stock and was grown overnight. The culture was diluted 1 : 100 into a 50 ml culture as described above. The culture was again incubated at 37 °C overnight, then centrifuged and the pellet washed with fresh LB medium to remove the β-lactamase produced during overnight incubation. Resuspended in the same volume (50 ml) of fresh LB medium containing ampicillin and chloramphenicol the culture was diluted 1 : 50 into either a 100 ml or a 1 L LB culture containing ampicillin and chloramphenicol. The culture was incubated at 37 °C until it reached an OD_{600} = 1, then cooled down to room temperature and induced with 1 mM IPTG. After induction it was incubated at 23 °C for 3 hrs. Finally, the culture was centrifuged at 2700 g and 4 °C for 15 min, the bacterial pellet frozen in liquid nitrogen and stored at –80 °C until further use.

4.7 Purification of recombinant RF-C from bacterial cell extracts

4.7.1 Buffer solutions

The following buffers were used for protein purification:

Lysis buffer A: 25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, 10 % (v/v) glycerol, 0.02 % (v/v) NP-40, 1 mM PMSF, 1 mM ATP and 1 μg/ml each of pepstatin / bestatin / leupeptin (from Bachem).

Lysis buffer B: 25 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 10 % (v/v) glycerol, 0.02 % (v/v) NP-40, 1 mM PMSF, 1 mM ATP and 1 μg/ml each of pepstatin / bestatin / leupeptin.

PC elution buffer 1: 25 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 5 mM $MgCl_2$, 10 % (v/v) glycerol, 0.02 % (v/v) NP-40, 1 mM PMSF, 1 mM ATP and 1 μg/ml each of pepstatin / bestatin / leupeptin.

PC elution buffer 2: 25 mM NaH_2PO_4 (pH 8.0), 600 mM NaCl, 5 mM $MgCl_2$, 10 % (v/v) glycerol, 0.02 % (v/v) NP-40, 1 mM PMSF, 1 mM ATP and 1 μg/ml each of pepstatin / bestatin / leupeptin.

PC dilution buffer: 25 mM NaH_2PO_4 (pH 8.0), 5 mM $MgCl_2$, 10 % (v/v) glycerol, 0.02 % (v/v) NP-40, 1 mM PMSF, 1 mM ATP and 1 μg/ml each of pepstatin / bestatin / leupeptin.

Nickel wash and elution buffers: 25 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 5 mM $MgCl_2$, 10 % (v/v) glycerol, 0.02 % (v/v) NP-40, 1 mM PMSF, 1 mM ATP and 1 μg/ml each of pepstatin / bestatin / leupeptin and an imidazole concentration varying from 5 mM to 500 mM. The precise concentration in the buffers is further described for each purification protocol (see “Results”).

4.7.2 Protein minipreparation: Nickel-NTA spin column

The bacterial pellet of a 50 ml or 100 ml culture was thawed and resuspended in 1 ml of either modified lysis buffer A (300 mM NaCl, no $MgCl_2$ and no ATP) or lysis buffer B. The cells were cracked by sonication at level 5, 50 % duty cycle. Six times six pulses were applied and in between the samples placed on ice to cool. Next, the lysed cells were centrifuged at 20,000 g and 4 °C for 30 min and the supernatant containing the soluble protein fraction loaded onto a nickel-NTA spin column (Qiagen) equilibrated with 600 μl modified lysis buffer A (300 mM NaCl, no $MgCl_2$ and no ATP) or lysis buffer B. The column was loaded in two steps by centrifugation at 50 g and 4 °C for 2 min. The column was washed with first 600 μl modified lysis buffer A (300 mM NaCl, no $MgCl_2$ and no ATP) or lysis buffer B, then with 600 μl of the same buffer supplemented with 5 mM imidazole. Bound protein was eluted from the column with three elutions of 100 μl modified lysis buffer A (300 mM NaCl, no $MgCl_2$ and no ATP) or lysis buffer B, both containing 250 mM imidazole. Between each wash / elution step the column was centrifuged at 50 g and 4 °C for 2 min.

4.7.3 Protein minipreparation: Nickel-IDA agarose beads

The bacterial pellet of a 100 ml culture was thawed and resuspended in 1 ml of lysis buffer B. Sonication and centrifugation were carried out as described in the nickel-NTA spin column protocol. The supernatant containing the soluble protein fraction was added to 50 µl of nickel-IDA agarose beads which had been equilibrated in lysis buffer B (ProBond™ Nickel Chelating Resin, from Invitrogen). The resulting slurry was incubated on a roller-shaker at 4 °C for 1 hr. The beads were washed twice, first with 600 µl lysis buffer B and then with 600 µl of the same buffer supplemented with 5 mM imidazole. Bound protein was eluted from the nickel-IDA agarose beads with two elutions of 100 µl modified lysis buffer B containing 250 mM imidazole. For each wash / elution step the slurry was again incubated at 4 °C on a roller-shaker for 10 min and then centrifuged at 50 g and 4 °C for 1 min.

4.7.4 Protein minipreparation: Phosphocellulose

The bacterial pellet of a 100 ml culture was thawed and resuspended in 1.8 ml lysis buffer A. The cells were cracked by sonication (level 5, 50 % duty cycle, 9 times 5 pulses) and then centrifuged at 20,000 g and 4 °C for 30 min. The supernatant containing the soluble protein fraction was added to 150 µl of phosphocellulose (P11, from Whatman) equilibrated in lysis buffer A. The resulting slurry was incubated on a roller-shaker at 4 °C for 2 hrs. It was then washed three times for 10 min, first with 2 ml, afterwards twice with 1 ml of lysis buffer A. Bound protein was eluted in a first step with two times 300 µl PC elution buffer 1 for 15 min. Finally, protein minipreparation was completed by a second elution step with three times 300 µl PC elution buffer 2 for 25 min. For each wash / elution step the phosphocellulose slurry was incubated at 4 °C on a roller-shaker and was then centrifuged at 50 g and 4 °C for 1 min.

4.7.5 Large scale RF-C purification: Optimised protocol for phosphocellulose / nickel combined batch-purification

The bacterial pellet of a 1 L culture was thawed and resupended in 20 ml lysis buffer A. The cells were cracked by sonication (level 6, 60 % duty cycle, 10 times 6 pulses, the cells were cooled down to 4 °C after 6 pulses) and then centrifuged at 17,000 g and 4 °C for 30 min. 5 ml of phosphocellulose were equilibrated in lysis buffer A and the supernatant containing the soluble protein fraction was added. The resulting slurry was incubated at 4 °C on a roller-shaker for 2 hrs. It was washed three times with 30 ml lysis buffer A for 10 min. Bound protein was eluted in a first step with two times 6 ml PC elution buffer 1 for 15 min. To complete the phosphocellulose purification step the proteins still bound to phosphocellulose were eluted with two times 5 ml PC elution buffer 2. Small samples of these last two fractions were put aside for further analyses. The rest of the two fractions was pooled, diluted 1 : 2 with

PC dilution buffer to a NaCl concentration of 300 mM and added to 200 µl nickel-IDA agarose beads (ProBond™ Nickel Chelating Resin, from Invitrogen) equilibrated in nickel wash buffer without imidazole. The resulting slurry was incubated at 4 °C on a roller-shaker for 2 hrs. The nickel-IDA agarose beads were then washed three times for 10 min, first with 4 ml then twice with 3 ml nickel wash buffer without imidazole. Bound protein was eluted from the beads with three step elutions where the imidazole concentration in the buffer was increased. The first nickel elution buffer contained 50 mM imidazole, the second 100 mM, and in the third nickel elution buffer the imidazole concentration was increased to 500 mM. For the 50 mM and the 100 mM imidazole step the beads were incubated twice in 200 µl nickel elution buffer containing the respective concentration of imidazole for 15 min. For the final 500 mM imidazole step the beads were incubated only once in 200 µl nickel elution buffer for 15 min. Throughout the purification procedure the fractions were always incubated at 4 °C on a roller-shaker and then centrifuged at 50 g and 4 °C for 1 min. Finally, 100 µl of each fraction were frozen in liquid nitrogen as 10 µl aliquots for RF-C activity assays.

4.7.6 Western blot analysis

The protein fractions of all purification procedures were analysed by SDS-PAGE and then electroblotted onto a nitrocellulose membrane (from Osmonics). The membrane was cut horizontally into two parts, whereas the upper part was incubated with a monoclonal mouse-α p140 antibody (the hybridoma cells were a gift from B. Stillman) diluted 1 : 10,000 in a TBST (containing 0.05 % Tween® 20) / 5 % (w/v) milk / 2 % (w/v) BSA solution. The lower part was treated with a mouse-α 6 x His antibody (from BAbCO) diluted 1 : 1000 in TBST / 3 % (w/v) BSA solution after blocking in the same protein solution without antibody. Alternatively, polyclonal antibodies against p40, p37 and p36 (rabbit-α p40, rabbit-α p37 and chicken-α p36, gifts from J. Hurwitz) were used in 1 : 2000 (α p37 and α p36) or 1 : 1000 (α p40) dilutions in a TBST / 2.5 % (w/v) milk solution after the membrane was blocked with TBST / 5 % (w/v) milk. After three washing steps with TBST solution the membranes were treated with horseradish peroxydase conjugated sheep-α mouse IgG (α p140 and α 6 x His), donkey-α rabbit IgG (α p40 and α p37), both from Amersham, or rabbit-α chicken IgY (α p36), from Promega. The membrane was again washed three times with TBST solution and treated with the ECL-system solutions (Enhanced Chemiluminescence, from Pierce). The bands corresponding to p140, p40, p37, p36 or the 6 x His-tagged subunits (Hisp37 and Hisp40) were visualised by exposing the membrane to an X-ray film.

4.7.7 Silver staining

To distinguish between the four small RF-C subunits the protein fraction was separated by SDS-PAGE. The gel was fixed in a solution containing 50 % (v/v) ethanol / 10 % (v/v) acetic acid in sterile, distilled water, then incubated in 30 % (v/v) ethanol for 15 min and washed three times for 5 min with sterile, distilled water. The gel was subsequently sensitised for 1.5 min in a solution containing 0.02 % (w/v) natriumthiosulphate pentahydrate ($Na_2S_2O_3 \cdot 5H_2O$). The solution was freshly prepared and a small volume kept for the developing solution. The gel was again washed three times for 30 s with sterile, distilled water. The gel was then stained with a 0.2 % (w/v) silvernitrate solution ($AgNO_3$) for 25 min at 4 °C and washed twice for 30 s with sterile, distilled water. Next, the silver stained gel was developed in a solution containing 6 % (w/v) sodiumcarbonate (Na_2CO_3), 0.05 % (v/v) formaldehyde (37 % solution) and 2 % (v/v) sensitising solution used before. The development process was stopped with a 6 % (v/v) acetic acid solution for 10 min when the desired level of staining was achieved, usually after 3 – 6 min. Finally, the gel was washed with sterile, distilled water at least four times for 15 min.

4.8 RF-C activity assays

4.8.1 DNA polymerase holoenzyme assay (on a circular DNA template)

A final reaction volume of 25 µl contained 120 ng of recombinant human PCNA, 500 ng of recombinant single-strand DNA binding protein (SSB), 10 u of recombinant human pol δ (Podust, et al., 2002, polymerase activity was determined in a standard poly dA / oligo dT polymerase assay) and 100 ng singly-primed pM13 DNA in a buffer containing 40 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT, 200 ng/µl BSA, 1 mM ATP and [^{3}H]dNTPs (40 µM dATP, 40 µM dGTP, 40 µM dCTP and 15 µM [^{3}H]dTTP which was labelled with [^{3}H] to 350 cpm/pmol, all purchased from Amersham). 1 µl of the RF-C fraction to be tested was added and the reaction incubated at 37 °C for 30 min. The reaction was stopped with 10 % (w/v) trichloroacetic acid (TCA) / 100 mM sodiumpyrophosphate (NaPP) for at least 20 min and then pipetted onto a GF-C filter (from Whatman) where large DNA molecules are retained. The filter was thoroughly washed with 100 mM NaPP / 3.2 % (v/v) HCl, dried with ethanol and soaked in scintillation fluid (Emulsifier-Safe™, from Packard Bio Science). The radioactive radiation emitted by the incorporated [^{3}H]-labelled nucleotides was measured in a scintillation counter.

4.8.2 DNA polymerase holoenzyme assay (on a linear DNA template)

A final reaction volume of 25 µl contained 120 ng of recombinant human PCNA, 500 ng of recombinant SSB, 10 u of recombinant human pol δ and 500 ng primed poly dA / oligo dT (1 : 20 ratio) in a buffer containing 40 mM Tris-HCl (pH 7.5), 7 mM $MgCl_2$, 1 mM DTT, 200 ng/µl BSA, 1 mM ATP and 20 µM [^{3}H]dTTPs which was labelled with [^{3}H] to 360 cpm/pmol. 1 µl of the RF-C fraction to be tested was added and the reaction incubated at 37 °C for 30 min. The samples were then analysed as described in chapter 4.8.1.

4.8.3 PCNA loading assay

Preparation of [γ^{32}P]-labelled PCNA: 500 ng phosphorylatable recombinant PCNA (ph-PCNA) was phosphorylated with 3 u of bovine heart muscle kinase (HMK, from Sigma) using 10 µCi of [γ^{32}P]-ATP in 20 µl of a buffer containing 20 mM Tris-HCl (pH 7.5), 1 mM DTT, 100 mM NaCl and 12 mM $MgCl_2$. The reaction was incubated at 37 °C for 20 min and then kept on ice until further use.

Loading assay: A final reaction volume of 25 µl contained 50 ng of [γ^{32}P]-ph-PCNA and 100 ng of pET21a plasmid DNA in a buffer containing 40 mM triethanolamine-HCl (pH 7.5), 0.2 mg/ml BSA, 1 mM DTT, 10 mM $MgCl_2$ and 1 mM ATP. 1 µl of the RF-C fraction to be tested was added and the reaction incubated at 37 °C for 3 min. The DNA-protein complexes were cross-linked for 10 min with 2.5 µl of 1 % (v/v) glutaraldehyde (freshly prepared from a 25 % stock solution). The samples were adjusted with loading buffer containing a final concentration of 2.5 % (w/v) Ficoll-400, bromphenol-blue and xylene cyanol and immediately loaded onto a 0.8 % agarose gel. The gel was run in 0.5 x TBE buffer containing 0.1 % (w/v) SDS, then fixed with 10 % (w/v) trichloroacetic acid (TCA) until the bromphenol-blue dye had turned green and finally soaked in 10 % (v/v) acetic acid / 12 % (v/v) methanol. The gel was dried under a vacuum without heating and finally autoradiographed.

4.8.4 Exo- / Endonuclease assay

A final reaction volume of 10 µl contained 200 ng double-stranded DNA (the PCR product (1052 bp) of the amplification of the p36 gene was used) in a buffer containing 40 mM Tris-HCl (pH 7.5), 1 mM DTT and 200 ng/µl BSA. 0.4 µl or 1 µl of the RF-C fraction to be tested was added and the reaction was incubated at 37 °C for 1 hr. As a nuclease positive control, 200 u of DNAse I (Böhringer, Mannheim) were added to the reaction mix instead of RF-C and this reaction was also incubated at 37 °C for 1 hr. Then the reaction mix was immediately adjusted with loading buffer for agarose gels (5 % (v/v) glycerol, 5 mM Tris-HCl (pH 7.5), 2.5 mM EDTA, 0.01 % (w/v) bromphenol-blue and 0.01 % (w/v) xylencyanol, final concentration) and loaded onto a 0.8 % agarose gel containing 0.5 µg/ml ethidiumbromide.

4.9 Supplier list for materials and devices

DNAs and nucleotides

dATP, dGTP, dTTP, dCTP	Amersham
[^{3}H]dTTP, [γ^{32}P]ATP	Amersham
oligo dT	Amersham
pET21a, pET41a	Novagen
poly dA	Amersham
primers	Microsynth
T7-promoter and terminator primers	Novagen

Bacterial strains

E. coli BL21(DE3)	Stratagene
E. coli BL21(DE3)-Codon Plus RIL	Stratagene
E. coli BL21(DE3)-Codon Plus RP	Stratagene
E. coli BL21(DE3)Gold	Stratagene
E. coli BL21(DE3)pLysS	Stratagene
E. coli DH5α	Invitrogen
E. coli XL1-Gold	Stratagene

Bacterial culture

Ampicillin, Kanamycin	Fluka
Bacto™ agar	Becton, Dickinson and Company
Chloramphenicol	Sigma
LB medium capsules	Q-BIOgene

Enzymes and reaction buffers

Ase I	New England Biolabs
BamH I	New England Biolabs
BSA (nuclease-free)	New England Biolabs
DNAse I	Roche
Eag I	New England Biolabs
Nde I	New England Biolabs
Pfu-polymerase	Promega
Pfu-polymerase reaction buffer	Promega
Reaction buffer NEB2/3/4 and NEBSal	New England Biolabs
Sal I	New England Biolabs
Spe I	New England Biolabs
T4 DNA ligase	New England Biolabs
T4 DNA ligase reaction buffer	New England Biolabs
Xba I	New England Biolabs
Xho I	New England Biolabs

DNA preparation kits

Nucleospin DNA purification kit	Macherey-Nagel
Qiagen buffers P1, P2, P3, PB, PE and QG	Qiagen
Qiagen Plasmid Midi and Maxi kit	Qiagen
QIAquick Gel Extraction kit	Qiagen
QIAquick PCR purification kit	Qiagen

Protein purification

Bradford assay solution	BioRad
Nickel-NTA spin column	Qiagen
P11 (phosphocellulose)	Whatman
Pepstatin, Bestatin and Leupeptin	Bachem
ProBond™ Nickel-Chelating Resin	Invitrogen

SDS-PAGE, western blot analysis and antibodies

Acrylamide 37.5 : 1	Serva
Chicken anti p36 polyclonal antibody	J. Hurwitz (gift)
ECL solutions	Pierce
High molecular range protein marker	Amersham
HRP conjugated donkey anti rabbit IgG	Amersham
HRP conjugated rabbit anti chicken IgY	Promega
HRP conjugated sheep anti mouse IgG	Amersham
Low molecular range protein marker	Amersham
Mouse anti 6 x His antibody	BAbCO
Mouse anti p140 monoclonal antibody	B. Stillman (gift)
Nitrocellulose membrane	Osmonics
Rabbit anti p37 polyclonal antibody	J. Hurwitz (gift)
Rabbit anti p40 polyclonal antibody	J. Hurwitz (gift)

Activity assays

Bromphenol blue	Sigma
Emulsifier-Safe™	Packard Bio Science
GF-C filter	Whatman
Heart muscle kinase	Sigma
PCNA and ph-PCNA	common stock of the Institute
Recombinant pol δ (Podust, et al., 2002)	E. Friedrich (gift)
Single-stranded pM13	common stock of the Institute
SSB	common stock of the Institute
Xylencyanol	Merck

All chemicals which are not listed here were purchased from Fluka.

Devices

Branson Sonifier cell disruptor 1315	Branson
Centrifuge Rotanta / P	Hettich
Cooling waterbath (16 °C)	Fryka Kältetechnik
Eppendorf centrifuge 5417C, rotor: F 45-30-11	Eppendorf
Eppendorf centrifuge 5417R (4 °C), rotor: F 45-30-11	Eppendorf
Gel electrophoresis / western blot apparatus	BioRad
Heat block	Techne Dri-block DB-2D
Incubators for bacterial cultures	Lab Therm, Kühner and Heraeus
Microwave	König
PCR machine	Perkin-Elmer
pH meter	Metrohm
Pipetman	Gilson
Power Supply Model 200 / 2.0	BioRad
Scales	Mettler
Scintillation counter MR 300	Kontron
Shaker / roller shaker	Heidolph / Greiner
Sorvall centrifuge RC5C, rotors: GS-3 and SS-34	Sorvall
UV-table (Reprostar)	CAMAG
UV-visible recording Spectrophotometer	Shimadzu
Vortex Genie	Bender & Hobein AG
Waterbath (37 °C)	GFL
White/UV transilluminator (gel documentation system)	UVP
X-ray film	Fuji
X-ray film cassette	Okamato
X-ray film processor FPM 100 A	Fuji

5 Results

5.1 Cloning strategy

Previous attempts to express human RF-C in a bacterial system had failed because of the fact that all five subunits were completely insoluble when induced individually. Therefore, the goal of this project was to express all five human RF-C subunits simultaneously. A key problem of the expression of multi-subunit complexes is that all subunits of the complex have to be induced on an equal level to facilitate complex formation. Since it is known for RF-C that the small subunits are only soluble when a complex of several subunits is formed, it was particularly important for the production of recombinant RF-C that an expression system was chosen where all five subunits could be induced equally. Therefore, the pET vector system was chosen. An additional advantage of this vector system is the fact that one can choose from a large selection of similar plasmids which only differ in the composition of the multiple cloning site or the antibiotic resistance. The multiple cloning sites of the more recently developed pET vectors, including pET21a and pET41a, contain sites for more than ten restriction enzymes, which greatly helped to design the cloning strategy for this multi-step cloning procedure. The restriction enzymes used during the cloning procedure had to be chosen very carefully to rule out the possibility that the enzymes digested the target genes. The goal was to construct first a set of two plasmids with different antibiotic resistance, one carrying the genetic information for p38 and p140, the other for Hisp37, Hisp40 and p36, and secondly a vector which had all five RF-C genes aligned, one after the other in the same orientation. Each gene carried its own ribosome binding site upstream of the start codon to ensure equal expression of all five genes. The cloning strategy is depicted in Figure 4 and described in detail below.

5.2 PCR amplification of the five human RF-C genes

Five different bacterial or yeast vectors, each containing the genetic information for one of the RF-C subunits, served as DNA templates for the PCR amplifications of the five RF-C genes. The genes of p40 and p37 already had the sequence for a 6 x His-tag added at the 5'-end of the genes in the original vector. Therefore the two genes were amplified including the information for the 6 x His-tag. Primers for each PCR amplification were designed as described in "Materials and Methods" according to the cloning strategy. Each primer consisted of a complementary region at the 3'-end, one or two sites for restriction enzyme digestion and six or eight bases (6 x A, 6 x G, 8 x A or 8 x G) at the 5'-end to enable restriction enzymes to cut the site close to the end of the DNA double-strand (see Table 6).

The PCR amplification mixes were prepared as described in "Materials and Methods". The PCR for the p140 gene amplification needed special optimisation to guarantee the synthesis

of the 3542 bp product with as few mutations as possible. This PCR reaction had to be run separately because of the different PCR cycle. The amplification of the genes of the four small subunits could be performed simultaneously. Samples of the amplified products were loaded onto a 1 % agarose gel containing 0.5 μg/ml ethidiumbromide to confirm the size of the expected PCR product and the level of amplification.

Figure 4 A Cloning strategy for human RF-C (1). For details see text.

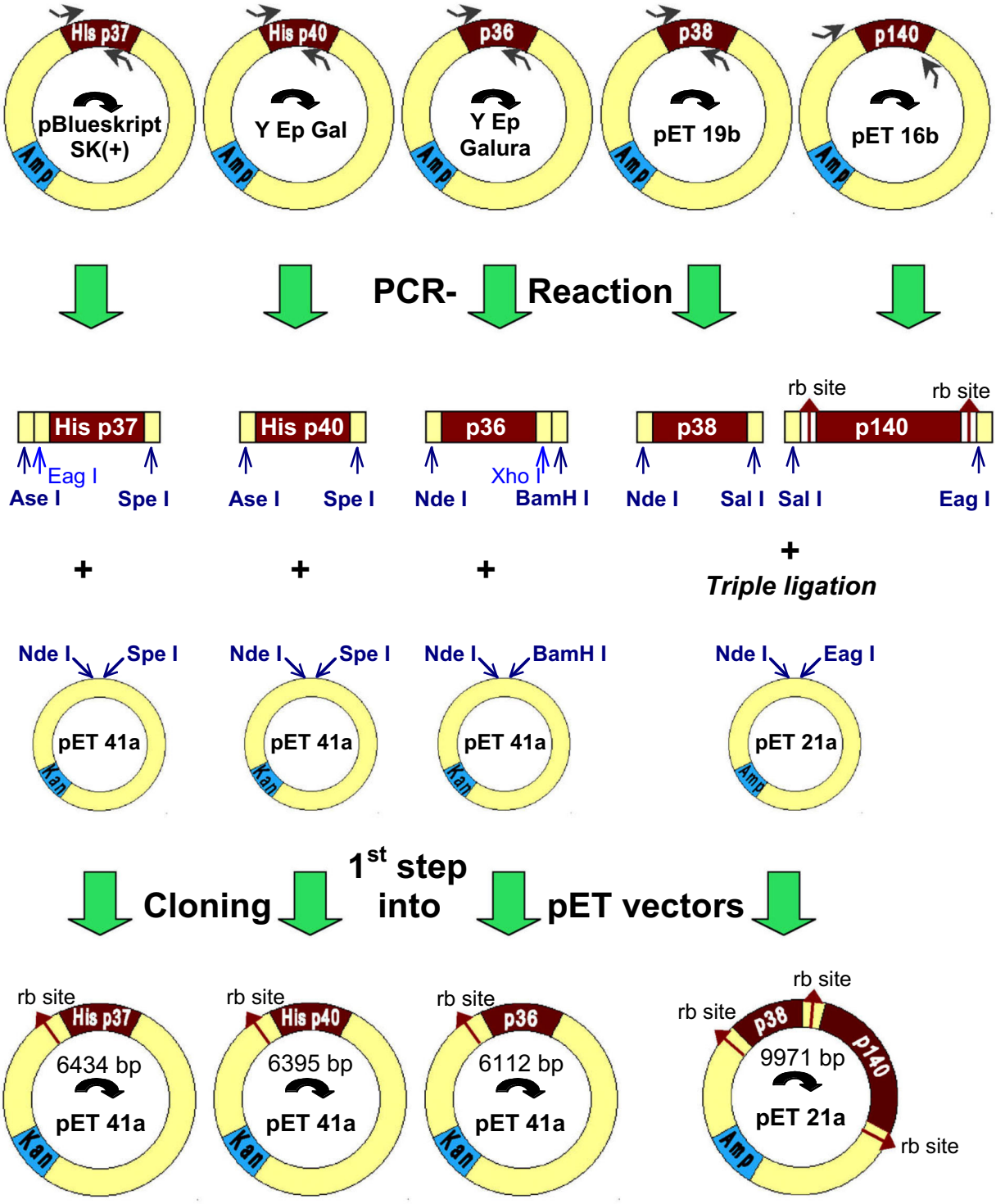
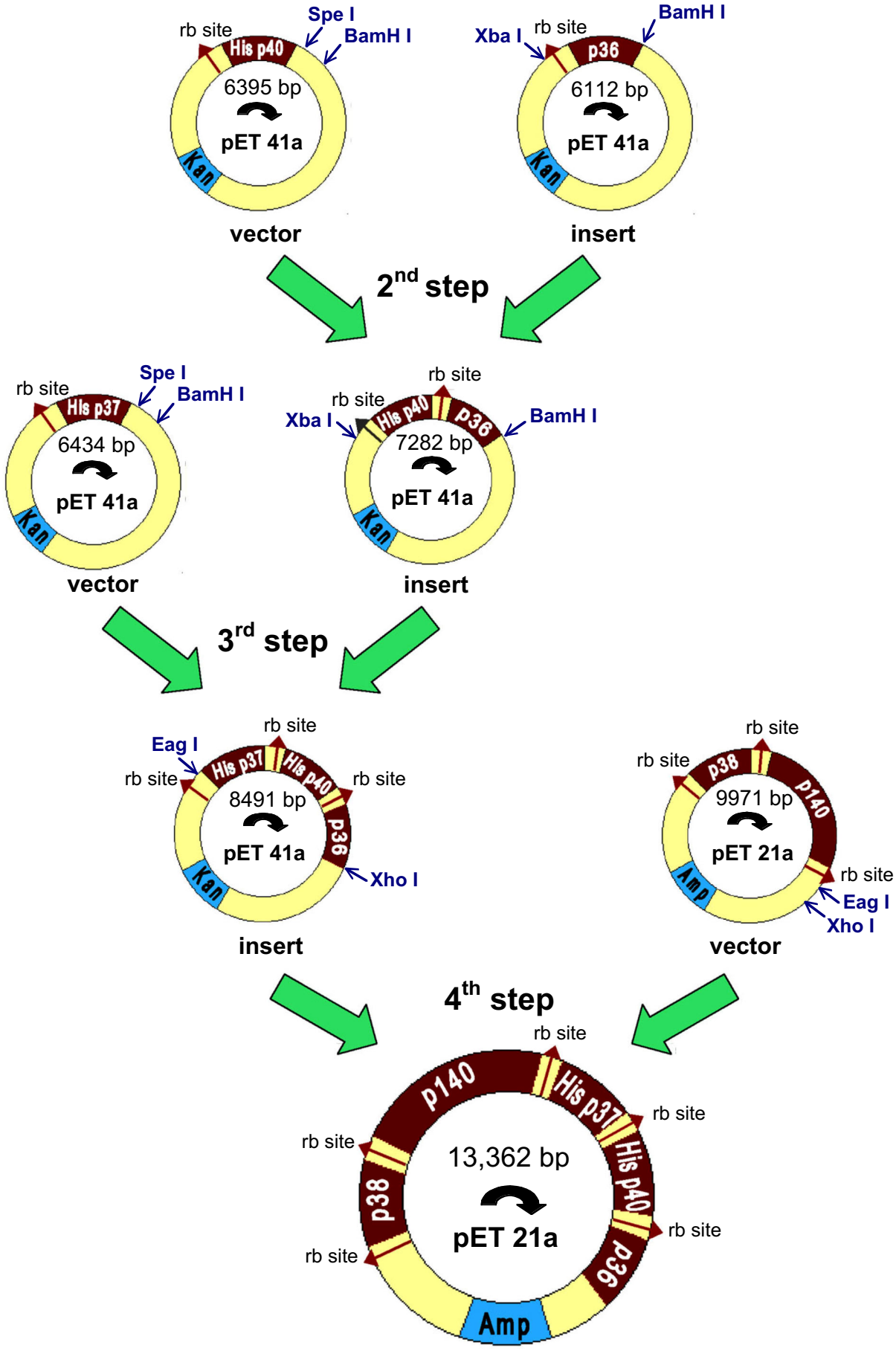


Figure 4 B Cloning strategy for human RF-C (2). For details see text.



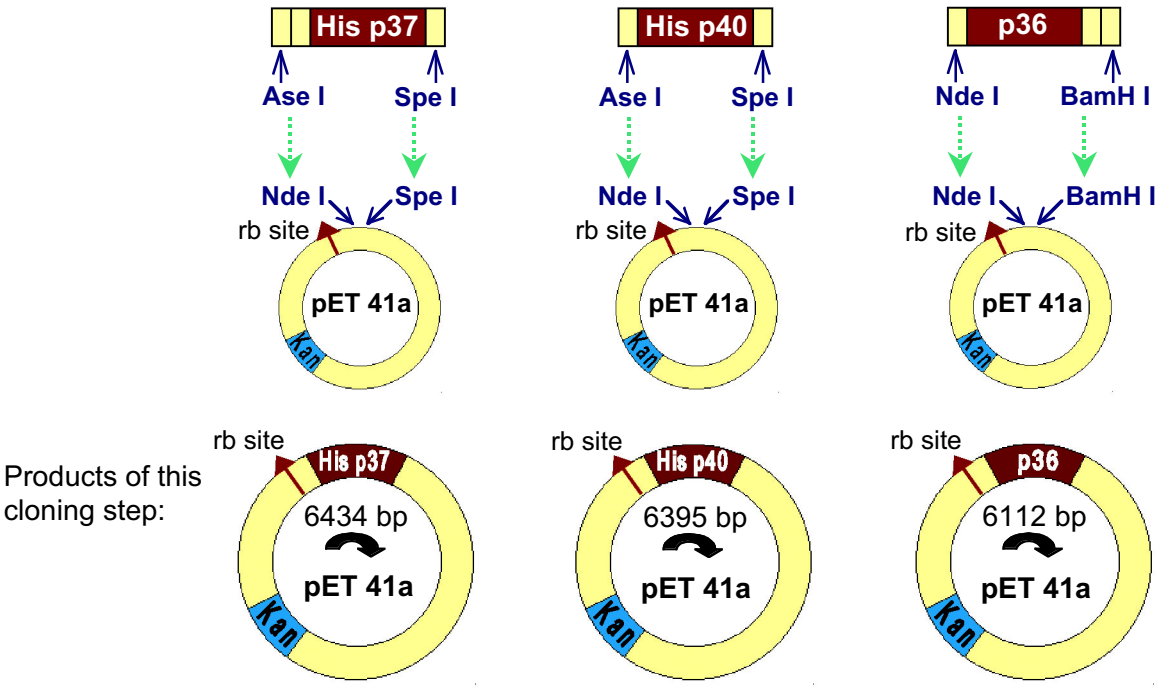
5.3 Cloning of the five RF-C genes into pET vectors (see overview in Figure 4 A)

5.3.1 Cloning of the Hisp37, Hisp40 and p36 genes into pET41a vector

For each cloning reaction 1 μg of pET41a vector DNA was digested with Nde I and Spe I (for Hisp37 and Hisp40 cloning) or Nde I and BamH I (for p36 cloning). All the reactions were incubated in buffer NEB2 containing BSA at 37 °C for 2.5 hrs. The PCR products of the Hisp37, Hisp40 and p36 genes were digested in two steps: The Hisp37 and Hisp40 DNAs were digested with Ase I (buffer NEB3), the p36 DNA with Nde I (buffer NEB4). The DNAs were purified by gel extraction and used for the second digestion step. The Hisp37 and Hisp40 genes were digested with Spe I (buffer NEB3, with BSA), the p36 gene with BamH I (buffer NEB2, with BSA). All the reactions were incubated at 37 °C overnight.

The double-digested vectors and genes were purified by gel extraction. 5 μl of each elution were analysed on a 0.8 % agarose gel and the insert : vector-ratio was calculated. For ligation the insert and vector DNA were mixed at a 2 : 1 ratio and the reaction was incubated as described in “Materials and Methods”. With this ligation the Nde I site in the pET41a/Hisp37 and pET41a/Hisp40 constructs was destroyed. Half of each ligation mix (10 μl) was used to transform 100 μl of competent *E. coli* DH5α cells. Half of this transformation mix was plated out on LB plates containing kanamycin. The plates were left at 37 °C overnight. Several colonies had grown on each plate: 40 colonies for pET41a/Hisp37, 50 colonies for pET41a/Hisp40 and 4 colonies for pET41a/p36. Five clones for Hisp37 and Hisp40 and all the p36 clones were selected, the plasmid DNA was purified by minipreparation and analysed by restriction enzyme digestion (data not shown). Two clones for each gene were sequenced.

Figure 5 Cloning of the Hisp37, Hisp40 and p36 genes into pET41a vectors

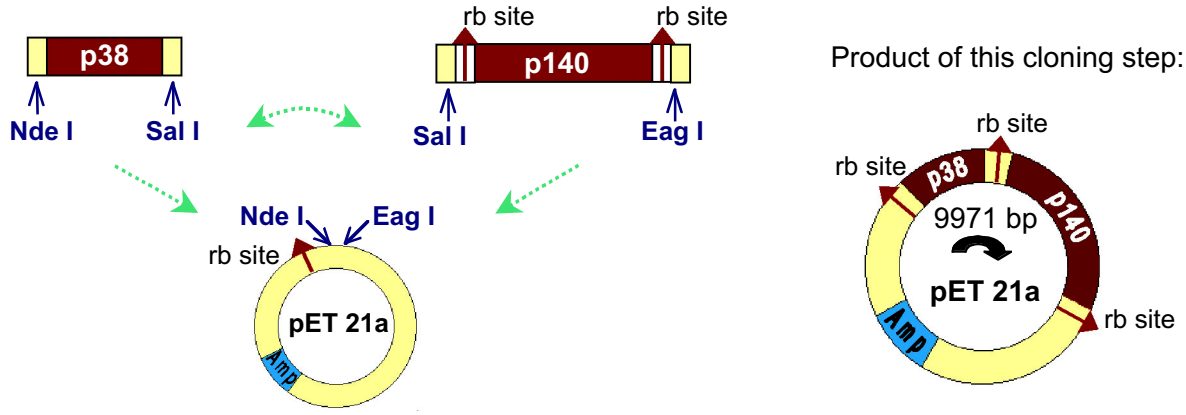


5.3.2 Cloning of the p38 and p140 genes into a pET21a vector

5 μg of pET21a vector DNA were digested with Nde I (buffer NEB4) as described in "Materials and Methods". The vector DNA was purified by gel extraction and digested with Eag I (buffer NEB3). The reaction mix was again incubated at 37 °C for 2.5 hrs. The PCR products of the p38 and p140 genes were digested with Sal I (buffer NEBSal I, with BSA). The reactions were incubated at 37 °C overnight. The DNAs were purified by gel extraction and then digested for the second time. The p38 gene was digested with Nde I (buffer NEB4), the p140 gene with Eag I (buffer NEB3). These reactions were again incubated at 37 °C overnight.

The double-digested vector and PCR product DNAs were purified by gel extraction. In this cloning step two genes (p38 and p140) were cloned into the pET21a vector by triple ligation. Therefore ligation mixes were prepared with different insert 1 : insert 2 : vector-ratios (p140 : p38 : pET21a). The following two ratios were used: 1 : 1 : 1 and 2 : 1.5 : 1. The ligation mixes were treated as described in "Materials and Methods". Half of each mix was used to transform 100 μl of competent *E. coli* DH5α cells. Half of either transformation mix was plated out on LB plates containing ampicillin and the plates were left at 37 °C overnight. Many colonies had grown on both of the plates. The plasmid DNA of a total number of 30 clones from both plates was purified by minipreparation. It was then loaded onto a 0.8 % agarose gel to check the size of the purified plasmid DNA (with two inserts: 9971 bp) compared to the pET21a vector (5443bp). 21 plasmid samples had the same size as the pET21a vector, but nine samples ran much higher on the gel (data not shown). To visualise the orientation of the two genes to the vector and to each other the plasmid DNA of these nine clones was digested with Nde I, which cut the vector three times (once just before the p38 gene and twice close to the 3'-end of the p140 gene). Five of the analysed nine clones showed the expected pattern on the gel (data not shown). Additionally, the remaining five clones were analysed by several PCR reactions using combinations of vector primers and primers which annealed to a sequence on the gene (see "Materials and Methods"). All five clones showed the expected PCR products on the agarose gel (data not shown) and two of them were sequenced.

Figure 6 Cloning of the p38 and p140 genes into a pET21a vector



5.4 Cloning of p36 into the pET41a/Hisp40 vector

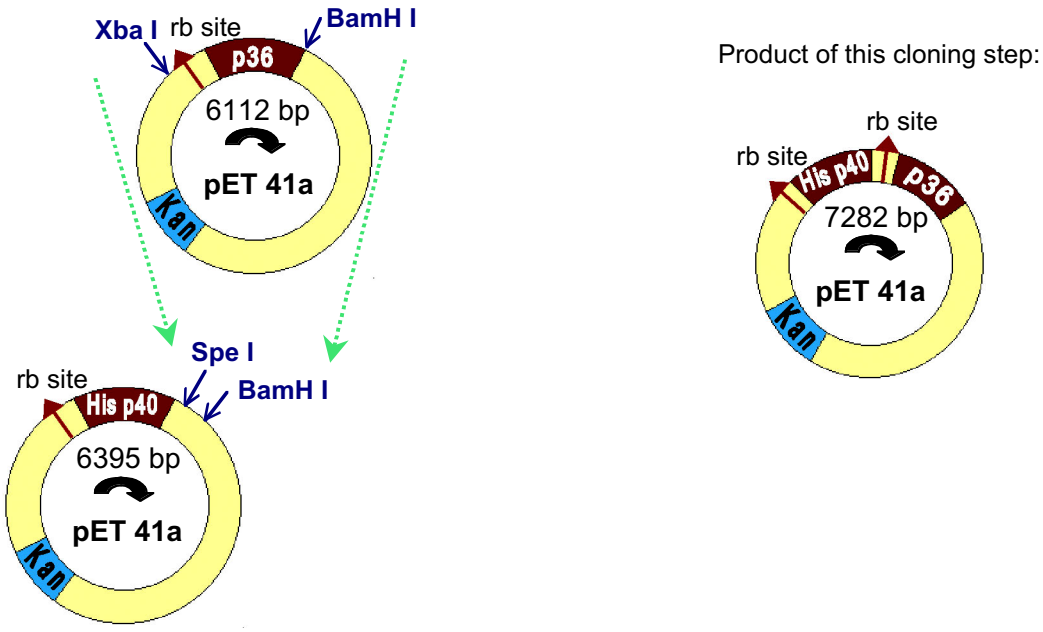
(see overview in Figure 4 B)

1 μg of pET41a/Hisp40 DNA was digested with Spe I and BamH I (buffer NEB2, with BSA). 1 μg of pET41a/p36 DNA was digested with Xba I and BamH I (buffer NEB2, with BSA). The reactions were incubated at 37 °C for 2.5 hrs. The whole reaction mixes were then loaded onto a 1 % preparative agarose gel. The bands corresponding to the linearised pET41a/ Hisp40 vector and the excised p36 gene were cut out of the gel under UV-light and the plasmid DNA was purified according to the gel extraction protocol.

Next, a sample of the purified DNA was analysed on a 0.8 % agarose gel and the insert : vector-ratio was calculated. For ligation insert and vector DNA were mixed at a 1 : 1 ratio and the reaction was incubated as described in “Materials and Methods”. Through ligation of the Xba I site on the insert to the Spe I site on the vector the site was destroyed for both enzymes. Half of the ligation mix was used to transform 100 μl of competent *E. coli* DH5α cells. Half of the transformation mix was plated out on an LB plate containing kanamycin. The plate was left at 37 °C overnight. 10 colonies had grown on the plate.

The plasmid DNA of five clones was purified by minipreparation. To confirm that the vector and the insert had ligated in the correct orientation a PCR reaction was designed using the upstream primer of the Hisp40 gene (pr 3, see Table 6) and the downstream primer of the p36 gene (pr 6, see Table 6), described in detail in “Materials and Methods”. The plasmid DNA of three clones was analysed, and all showed a PCR product of the correct size (data not shown).

Figure 7 Cloning of p36 into the pET41a/Hisp40 vector

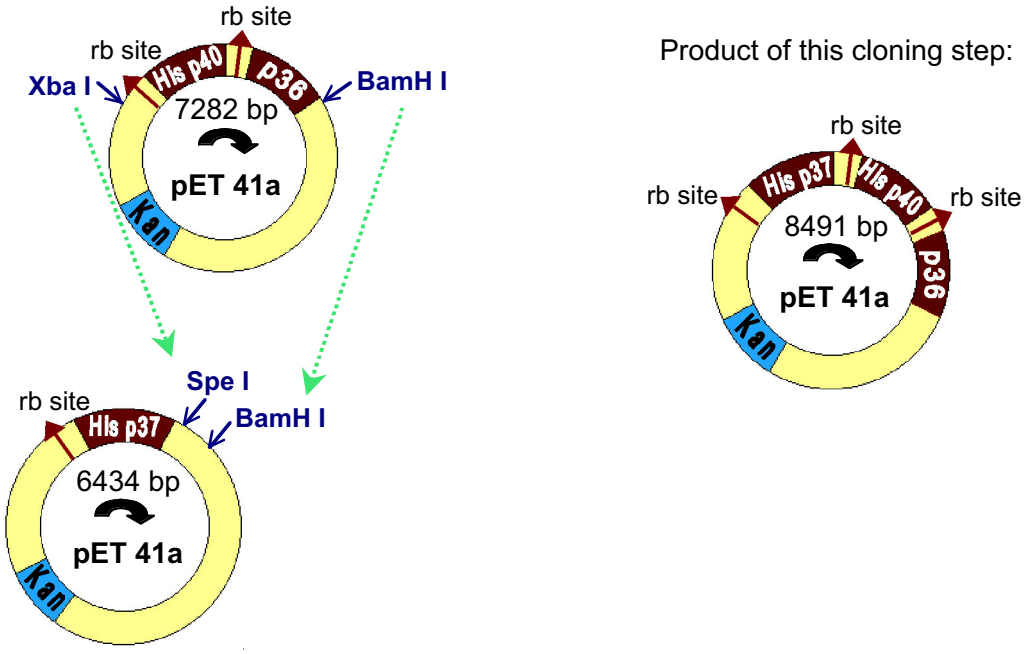


5.5 Cloning of Hisp40/p36 into the pET41a/Hisp37 vector

(see overview in Figure 4 B)

6 µg of pET41a/Hisp40/p36 DNA were digested with Xba I and BamH I (buffer NEB2, with BSA). The same amount of pET41a/Hisp37 DNA was digested with Spe I and BamH I (buffer NEB2, with BSA). The reaction mixes were incubated at 37 °C for 3 hrs and then loaded onto a 1 % preparative agarose gel. The bands corresponding to the linearised pET41a/Hisp37 vector and the excised Hisp40/p36 genes were cut out of the gel under UV-light and the plasmid DNA was purified according to the gel extraction protocol. The purified DNA was again analysed on a 0.8 % agarose gel and the insert : vector-ratio was calculated. For ligation insert and vector DNA were mixed at two different ratios (1.5 : 1 and 2 : 1) and the reaction was incubated as described in "Materials and Methods". By ligating the Spe I site on the vector to the Xba I site on the insert the site was destroyed for both enzymes. Then half of each mix was used to transform 100 µl of competent *E. coli* DH5α cells. Half of the transformation mix was plated out on LB plates containing kanamycin. The plates were left at 37 °C overnight. On one plate (1.5 : 1 ratio) eight colonies had grown, whereas on the other (2 : 1 ratio) only five colonies had grown. The plasmid DNA of four clones from each plate was purified by minipreparation. To verify and visualise the correct ligation the purified plasmid DNA was analysed by PCR using a primer pair for the Hisp37 gene (pr 1 and pr 2, see Table 6) and a primer pair for the Hisp40 gene (pr 3 and pr 4, see Table 6), as described in "Materials and Methods". The analysed clones showed PCR products of the correct size on the gel (data not shown).

Figure 8 Cloning of Hisp40/p36 into the pET41a/Hisp37 vector



5.6 Cloning of Hisp37/Hisp40/p36 into the pET21a/p38/p140 vector

(see overview in Figure 4 B)

5 μg of pET21a/p38/p140 DNA and 5 μg of pET41a/Hisp37/Hisp40/p36 DNA were digested in two steps. Both plasmids were first linearised with Eag I. The reactions were incubated in buffer NEB3 at 37 °C for 4.5 hrs. The DNA was purified by gel extraction and was again digested at 37 °C for 4.5 hrs with Xho I (buffer NEB2, with BSA). The whole reaction mixes were then loaded onto a 0.8 % preparative agarose gel. The bands corresponding to the linearised pET21a/p38/p140 vector and to the excised Hisp37/Hisp40/p36 insert were cut out of the gel. The DNA was purified according to the gel extraction protocol. A sample of the digested and purified DNA was analysed on a 0.8 % agarose gel and the insert : vector-ratio was calculated. For ligation insert and vector DNA were mixed at two different ratios (1 : 1 and 2 : 1) and the reaction was incubated as described in “Materials and Methods”. Half of each mix was used to transform 100 μl of competent *E. coli* DH5α cells and *E. coli* XL10-Gold, respectively. The transformation mixes were centrifuged at 100 g for 5 min, resuspended in 100 μl LB medium and plated out on LB plates containing ampicillin. The plates were left at 37 °C overnight. Only very few colonies had grown on some of the plates and most of them were very small. The four largest colonies were picked and the plasmid DNA was purified by minipreparation. The purified plasmid DNAs were analysed on a 0.8 % agarose gel to compare their size with the size of the pET21a/p38/p140 vector (9971 bp) and the pET41a/Hisp37/Hisp40/p36 vector (8491 bp). One clone (deriving from the 1 : 1 ratio ligation, *E. coli* DH5α cells) showed bands corresponding to a very large plasmid (pET21a/p38/p140/Hisp37/Hisp40/p36: 13,362 bp) (data not shown). This clone was further analysed by PCR using several combinations of vector primers and gene-specific primers. To detect each of the five RF-C genes separately, five PCR reactions were performed. The following primers were used (see Table 6). The results of this experiment are shown in Figure 9 C:

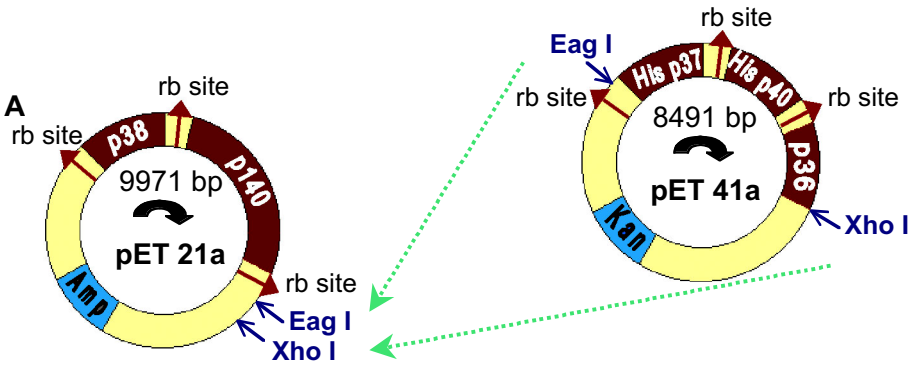
p38:	T7-promoter pr. + pr 8	A vector primer for this first gene in combination with a gene-specific primer visualised the correct orientation of the five gene construct (PCR product: 1178 bp).
p140:	pr 9 + pr 10	Gene-specific primers which anneal to flanking sites on both sides of the gene (PCR product: 3542 bp). These flanking sites are very similar in all five genes, therefore side products of the size of the shorter genes were also synthesised.
Hisp37:	pr 1 + pr 2	Gene-specific primers (PCR product: 1188 bp)
Hisp40:	pr 3 + pr 4	Gene-specific primers (PCR product: 1143 bp)
p36:	T7-terminator pr. +pr 5	A gene-specific primer for this last gene in combination with a vector primer visualised the correct orientation of the five gene construct (PCR product: 1173 bp).

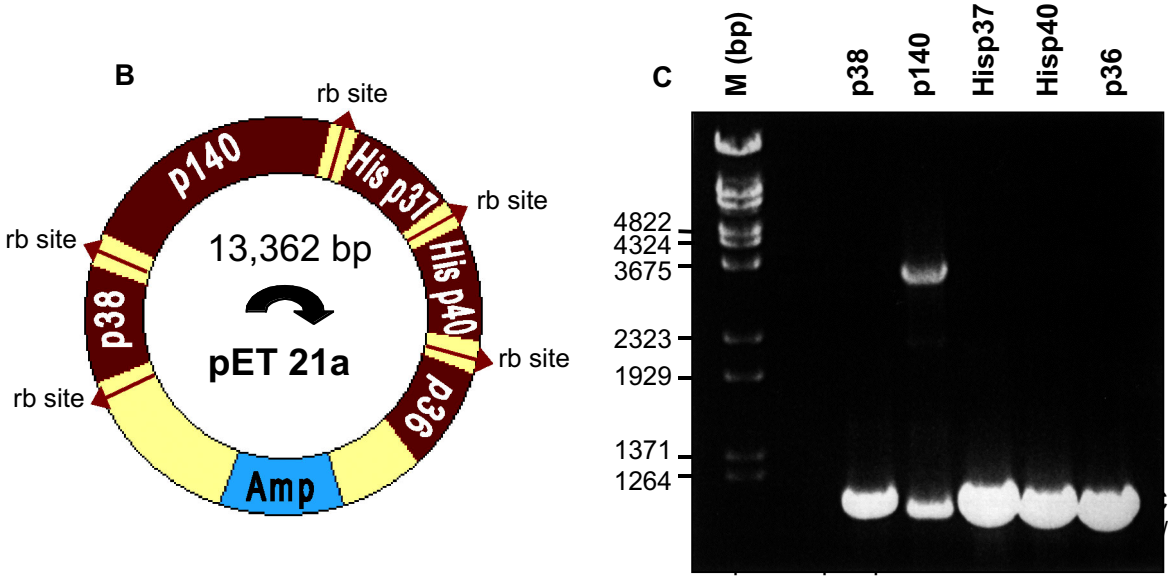
Figure 9 Cloning of Hisp37/Hisp40/p36 into the pET21a/p38/p140 vector

A: Cloning strategy

B: Product of this cloning step

C: PCR analysis of the vector construct containing all five RF-C genes





5.7 Expression

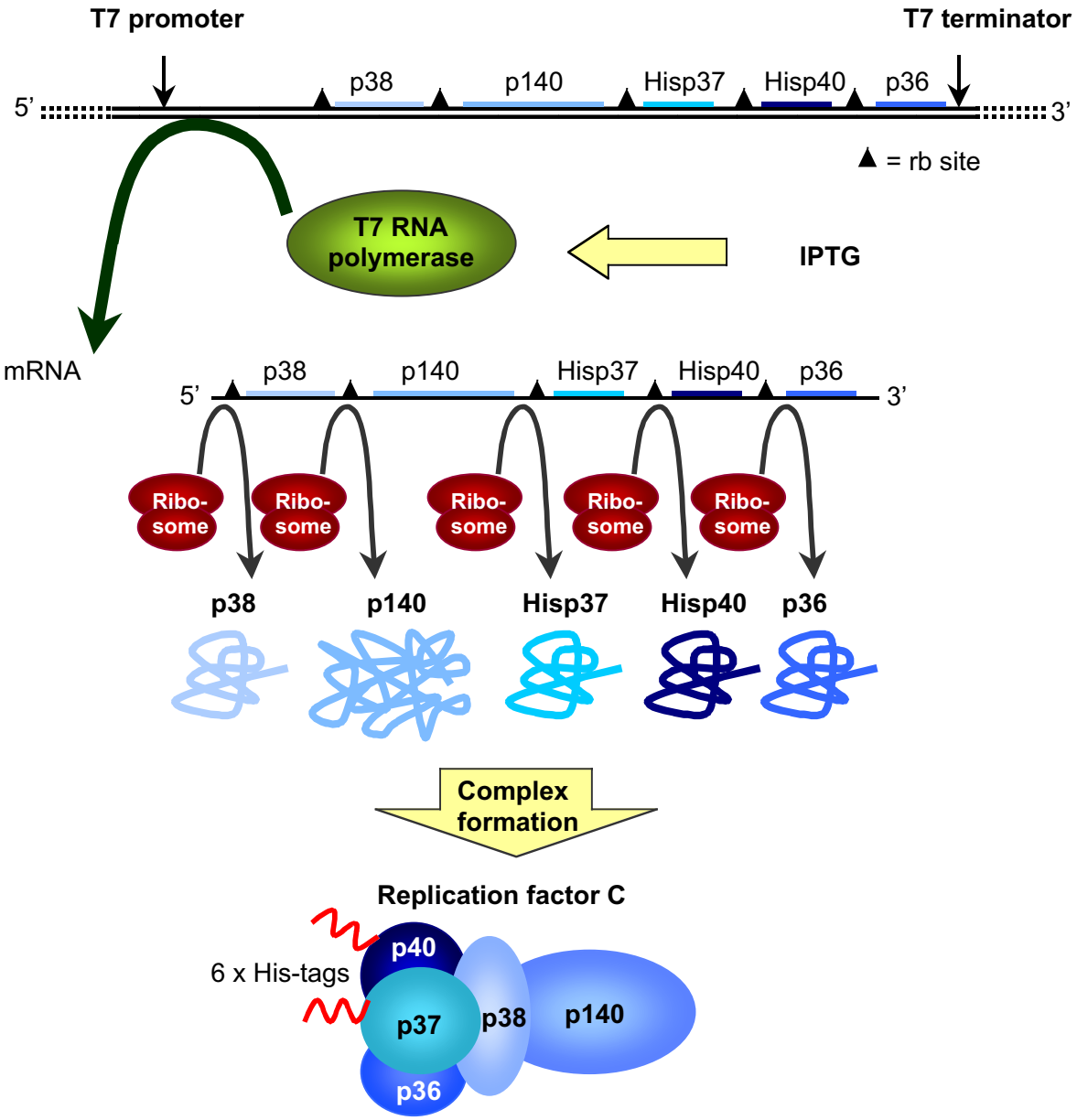
Several points have to be considered when an *E. coli* system is chosen for the expression of heterologous proteins to make optimal use of the advantages of the system. Therefore the method used for the overexpression of recombinant RF-C is described in detail below and in Figure 10.

For tight control of the expression of the five RF-C subunits the pET expression system was chosen. The target genes, the five human RF-C subunits, were cloned into a pET21a plasmid as described in the previous chapters. In this pET vector the genes to be expressed can be cloned under the control of the strong bacteriophage T7 promoter. The pET vector system was also chosen because it shows a very low level of basal target gene expression before induction. Even a low basal expression level of a toxic target gene could completely prevent bacterial cell growth or even cause cell death. Since human RF-C shows considerable sequence similarity with the *E. coli* γ-clamp loader, it might interfere with the bacterial life cycle and therefore be toxic for *E. coli* cells.

Transcription was induced by providing a source for the T7 RNA polymerase in the host cell. This was achieved by transforming *E. coli* strains which contained a chromosomal copy of the T7 RNA polymerase gene under *lacUV5* control (later referred to with the suffix DE3), such as the *E. coli* BL21 strain and its various derivatives. Therefore the T7 RNA polymerase was induced by adding IPTG to the bacterial culture. As a result a messenger RNA (mRNA) of approximately 8150 bases was transcribed, containing the information for the translation of all five RF-C subunits.

Translation was performed on the bacterial ribosomes. For efficient and even translation of the different subunits each RF-C gene carried its own ribosome binding site (rb site) upstream of the start codon. Equal expression of all five subunits was particularly important for the overexpression of RF-C, because only the five subunit RF-C complex is biochemically active. Additionally, it had been shown that singly overexpressed RF-C subunits were completely insoluble (Chen, et al., 1992a, Chen, et al., 1992b, Luckow, et al., 1994). In order to solubilise overexpressed RF-C in an enzymatically active form the formation of the five subunit complex had to be facilitated.

Figure 10 Transcription and translation of the pET21a/p38/p140/Hisp37/Hisp40/p36 construct



Transcription of the five RF-C genes is induced by addition of IPTG to the bacterial culture. Translation of the mRNA is performed on the bacterial ribosomes. Finally, the five RF-C subunits form a pentameric clamp loader complex. For details see text.

There are some well-known obstacles in the expression of difficult target proteins. One is the frequency of codon usage, which is different in the expression host, in this case *E. coli*, and the organism from which the genes to be expressed were cloned. As shown in "Problem and Approach" (Table 5) the frequency of codon usage markedly varies between the *E. coli* genome and the RF-C genes, even between the genes of the different RF-C subunits. Upon induction of the RF-C genes, which were cloned under the control of the strong T7 promoter, virtually all resources of the bacterial cell are used for the expression of the target genes. This might lead to a depletion of rare tRNAs in the host cell and to early termination of trans-

lation or misincorporation of amino-acids both resulting in degradation, loss of enzyme activity or even insolubility (formation of inclusion bodies). The problematic rare codons and the respective tRNAs are well characterised in *E. coli*. They include the arginine codon (argU: AGA or AGG), the isoleucine codon (ileY: ATA), the leucine codon (leuW: CTA) and finally the proline codon (proL: CCC). To circumvent this so-called codon-bias new *E. coli* expression strains were developed which carry a plasmid with the genetic information for the overexpression of certain rare tRNAs. They derive from an *E. coli* BL21 strain which lacks the Lon protease and the OmpT protease. The *E. coli* BL21(DE3)-Codon Plus RIL strain carries extra copies of the genes coding for tRNAs which recognise the arginine codons AGA and AGG, the isoleucine codon ATA and the leucine codon CTA. On the other hand, the *E. coli* BL21(DE3)-Codon Plus RP strain overexpresses the tRNAs which recognise the arginine codons AGA and AGG and the proline codon CCC.

For the overexpression of RF-C the *E. coli* BL21(DE3)-Codon Plus RIL strain was the first choice because the arginine codons (AGA and AGG) were critical in all RF-C subunits and the isoleucine codon (ATA) was critical in several subunits, including the large subunit of RF-C. The proline codon was also critical in three of the small subunits, but not in p140. Therefore both strains had to be compared with an ordinary *E. coli* BL21(DE3) strain as described later.

Early induction experiments with an *E. coli* BL21(DE3)-Codon Plus RIL strain expressing RF-C from the pET21a/p38/p140/Hisp37/Hisp40/p36 plasmid according to a standard induction protocol showed, that only a poor expression level was achieved and RF-C was almost completely insoluble (data not shown).

Every step of expression had therefore to be optimised individually to increase the expression level and solubility. An additional difficulty appeared during the optimisation procedure: The large subunit of RF-C, p140, always showed a high degree of degradation. Therefore, the optimisation procedure for the overexpression of RF-C had three well-defined goals:

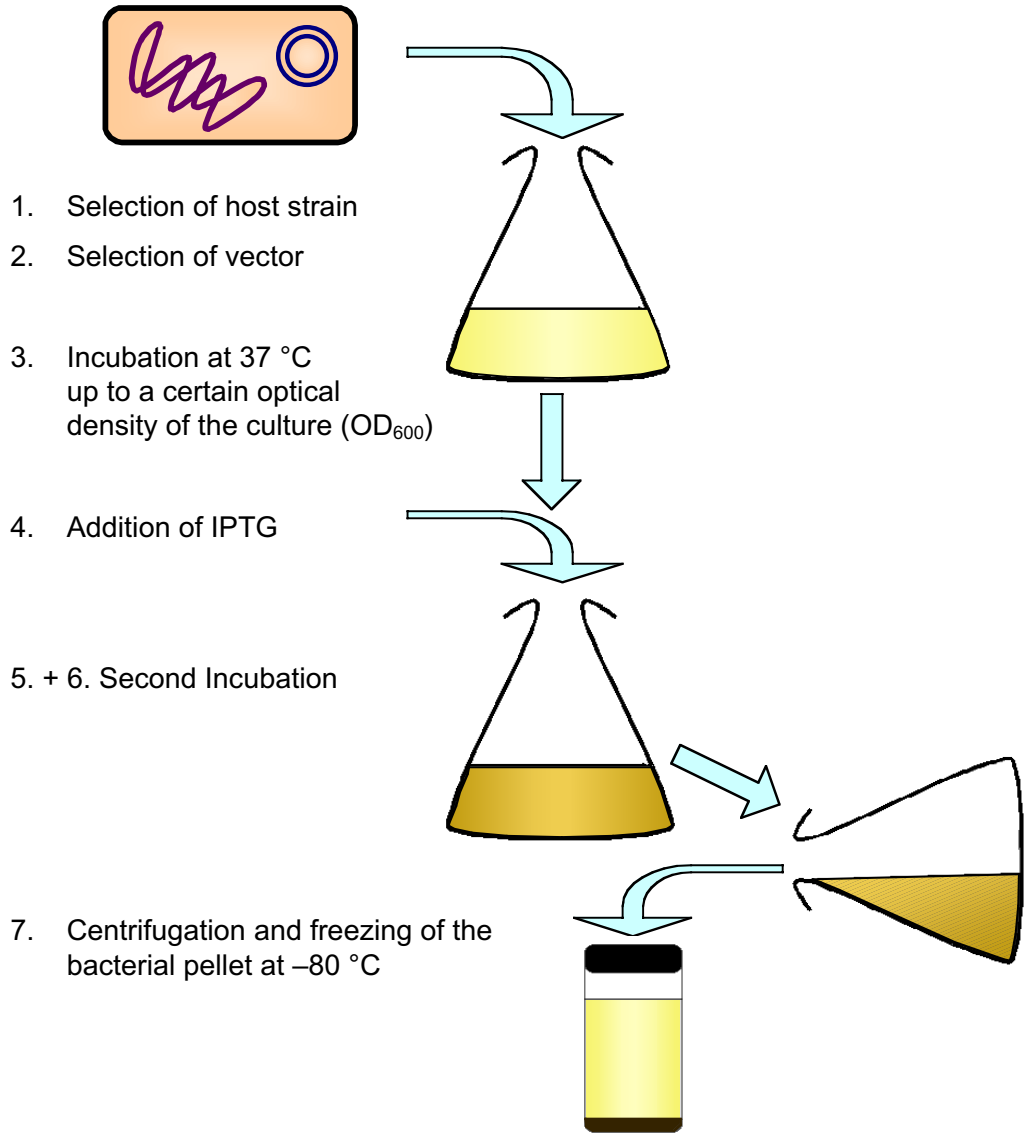
1. Maximal expression level of all subunits
2. Maximal solubility of the RF-C complex
3. Minimal degradation of p140

It was soon clear that optimisation towards a higher expression level led to less solubility, and higher solubility led to more degradation of the p140 subunit. The three goals were contradictory in many respects, the golden mean had to be found where soluble RF-C was expressed to a reasonable level and degradation of p140 was prevented.

Several parameters of a bacterial expression system can be varied and were optimised one by one. The host strains, plasmid vectors and the induction conditions listed below were compared in numerous small scale induction tests. Figure 11 shows an overview of the expression procedure.

1. Host strain	BL21(DE3), BL21(DE3)Gold, BL21(DE3)-Codon Plus RIL, BL21(DE3)-Codon Plus RP, BL21(DE3)pLysS
2. Expression vector for the RF-C genes	pET21a/p38/p140/Hisp37/Hisp40/p36 pET21a/p38/p140 + pET41a/Hisp37/Hisp40/p36
3. Optical density of the culture at time of induction	OD_{600} = 0.6, OD_{600} = 1
4. Concentration of IPTG	0.5 mM, 1 mM
5. Incubation temperature (after induction)	37 °C, 30 °C, 23 °C
6. Incubation time (after induction)	1 hr, 2 hrs, 3 hrs, 4 hrs

Figure 11 Induction of target gene expression



First, the above mentioned BL21(DE3) expression strains were compared in standard small scale induction tests using the pET21a/p38/p140/Hisp37/Hisp40/p36 vector and the induction conditions recommended in the manufacturer's protocol (see "Materials and Methods"). For each strain the total protein in bacterial cell extracts of samples taken before induction and 3 hrs after induction was analysed by SDS-PAGE and Coomassie-blue staining of the SDS-polyacrylamide gel to determine the expression level of RF-C. Additionally, the 3 hrs-after-induction samples were divided into soluble and insoluble protein fraction by centrifugation and were also loaded onto the SDS-polyacrylamide gel. None of the strains showed a high expression level of RF-C, indeed on the Coomassie-blue stained gel the individual subunits could not even be localised. The analysis of the BL21(DE3)-Codon Plus RIL strain was of particular interest, but in these tests neither the expression level nor solubility of RF-C showed any differences to the other strains tested (data not shown). Nevertheless, the two Codon-Plus strains were used in parallel for further optimisation steps, since codon-bias might become an issue when the expression level was markedly increased.

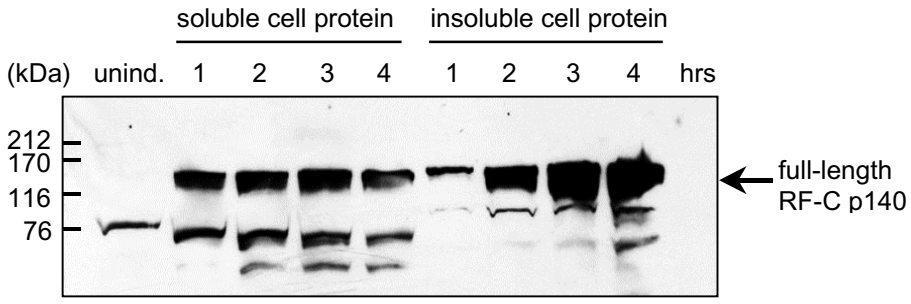
Next, the vectors from which the five subunits were expressed were compared. The three *E. coli* strains BL21(DE3) Gold, BL21(DE3)-Codon Plus RIL and BL21(DE3)-Codon Plus RP were doubly transformed with the two plasmids pET21a/p38/p140 and pET41a/Hisp37/Hisp40/p36. To prevent the loss of any of the plasmids, these strains had to be cultured in medium containing ampicillin (pET21a), kanamycin (pET41a) and chloramphenicol (RIL / RP plasmid) in the case of the Codon Plus strains. They were compared with the same host strains carrying the pET21a/p38/p140/Hisp37/Hisp40/p36 plasmid construct, which had been used in the previous experiment. For each host / construct combination a small scale standard induction test was run and the samples were analysed as described in the last experiment. Again, no difference in the expression level or in solubility was found between the host / vector combinations (data not shown).

The parameters concerning induction conditions, OD_{600} of the culture at the time of induction, the IPTG concentration and the incubation temperature after induction, were optimised together. Since overexpressed RF-C might be toxic for the expression host or interfere with its life cycle, the bacterial culture was grown to a very high optical density before induction (OD_{600} = 1). An additional advantage was the fact that at such a high OD_{600} more expression host cells per culture volume were available for target gene expression, which increased the expression rate and shortened the period of time needed to achieve the maximal expression level. First, all combinations of IPTG concentration 0.5 mM or 1 mM and incubation temperature after induction of 30 °C or 37 °C were tested with small scale induction tests. Samples were taken 3 hrs after induction, and the soluble and insoluble protein fractions analysed by SDS-PAGE and subsequent western blotting against the p140 subunit of RF-C. The

results showed that an increase of the induction temperature from 30 °C to 37 °C markedly increased the expression level of p140, but also lead to more degradation of this subunit. But neither of two induction temperatures tested had an influence on the solubility of expressed RF-C and all full-length RF-C p140 was completely insoluble. A high concentration of IPTG also increased the expression level, but hardly affected p140 degradation (data not shown).

In a final attempt to produce full-length and soluble RF-C a culture was induced with 1 mM IPTG and was then incubated at 23 °C using a cooling incubator for 4 hrs. Samples of the culture were taken at the time of induction and then every full hour after induction. The soluble and insoluble protein fractions of each sample were separated on an SDS-polyacrylamide gel and analysed by western blotting against the p140 subunit of RF-C (Figure 12). At last full-length p140 was detectable in the soluble protein fractions and considerably less degradation was observed. The expression time course also showed that the soluble full-length p140 was mainly produced during the first hour and increased only slightly during the second and third hour after induction, whereas in the insoluble protein fractions the amount of p140 was still on the increase after 4 hrs of induction. It was therefore concluded that a 3 hr induction period was optimal for the overexpression of RF-C.

Figure 12 Western blot analysis of the optimised RF-C expression and solubility test



Induction time course when the culture was induced at 23 °C. An anti-p140 antibody was used for this western blot analysis.

To summarise the optimal induction conditions for the overexpression of recombinant RF-C in an *E. coli* expression system: An *E. coli* BL21(DE3)-Codon Plus RIL strain was chosen to express recombinant RF-C from a pET21a vector carrying the genetic information for all five RF-C subunits. For induction the bacterial culture was grown at 37 °C in LB medium containing ampicillin and chloramphenicol to a high optical density (OD_{600} = 1). Target gene expression was then induced with a high concentration of IPTG (1 mM). The culture had to be cooled down to room temperature before addition of IPTG and was afterwards incubated at 23 °C in a cooling incubator for 3 hrs. Lowering the induction temperature to 23 °C dramatically increased the amount of soluble and full-length RF-C p140 (Figure 12).

5.8 Purification of recombinant RF-C from bacterial cell extract

5.8.1 On the topic of the 6 x His-tags added to two RF-C subunits

To purify RF-C from mammalian cell extracts several column purification steps are necessary and RF-C activity often is lost before homogeneity of the protein complex is achieved. Recombinant proteins are often expressed as tagged proteins. Tags can be added to the N- or the C-terminus of a protein. They are used as labels for protein detection and to distinguish between recombinant and endogenous protein. In this case the tagged protein is recognised by a specific antibody raised against the tag. The binding properties of the tag to a specific binding partner can be used for target protein purification. These interactions include enzyme-substrate interactions (e.g. in the case of the glutathione S-transferase-tag binding to glutathione) and chelator-ion interactions (in the case of the 6 x His-tag, which can chelate bivalent metal ions such as nickel). When proteins like glutathione S-transferase are fused to a target protein, the tag makes out a considerable part of the expressed fusion-protein (size of the glutathione S-transferase-tag: 26 kDa). The tag might therefore interfere sterically with protein-protein interactions involving the target protein, including complex formation, or inhibit enzymatic activities of the recombinant protein. Since both issues, complex formation and enzyme activity, are absolutely essential in the case of recombinant RF-C, a small tag was preferred to label overexpressed RF-C. A 6 x His-tag, where six histidine amino-acids (size of the 6 x His-tag: 0.84 kDa) are added to either the N- or the C-terminal end of the target protein, suited these purposes better than any other tag.

To define the RF-C subunits to be tagged and the location of the tag in each subunit the literature was searched for information about the three dimensional structure of the RF-C complex and the function of each subunit in the complex (reviewed in Jeruzalmi, et al., 2002 and Mossi and Hubscher, 1998). As already mentioned in the introduction the complex formation domain is located at the C-terminal end of all RF-C subunits where the amino-acid sequence is unique to each subunit (see Figure 1). When 6 x His-tags were placed at the N-termini of the subunits they were very likely located on the outside of the RF-C complex and were therefore less likely to interfere with complex formation. Two of the small RF-C subunits were selected as 6 x His-tag carriers; RF-C p37 and RF-C p40. The large RF-C subunit p140 and RF-C p38 were unsuitable for tagging. The large RF-C subunit is the main catalytic subunit and even a small tag might interfere with the enzymatic activity of the recombinant RF-C complex. RF-C p38 was found to have a linker function between the RF-C p40-p37-p36 subcomplex and p140 and hence contains two regions for complex formation, one at the N- and one at the C-terminus of the protein.

5.8.2 Protein minipreparation: Nickel spin column and nickel beads

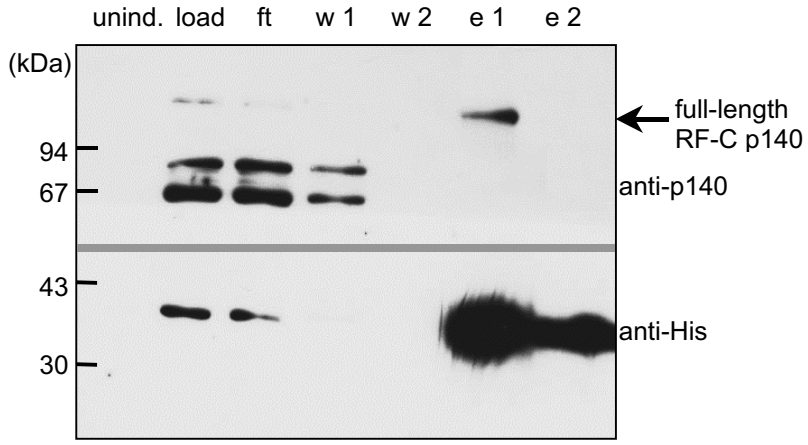
Since overexpressed RF-C contains two N-terminally 6 x His-tagged subunits, Hisp40 and Hisp37, its ability to bind to a nickel-NTA resin was tested in initial protein minipreparation experiments. The bacterial pellet of a 40 ml culture, where RF-C was expressed in BL21(DE3)-Codon Plus RIL cells either from the pET21a/p140/p38/Hisp37/Hisp40/p36 construct or from the combination of the pET21a/p140/p38 and pET41a/Hisp37/Hisp40/p36 plasmids, was prepared as described in “Materials and Methods” and the fraction containing the soluble cell protein was loaded by centrifugation onto a commercially available nickel-NTA spin column. The spin columns were washed and the bound protein was eluted with 250 mM imidazole in the elution buffers. All protein fractions were analysed by SDS-PAGE and subsequent western blotting against RF-C p140 and the 6 x His-tag. Equal signals for p140 and 6 x His-tagged subunits were detectable in the load and the flow through fractions, but no p140 and only a weak His-signal was found in the elution fractions (data not shown). This result did not depend on the vector(s) used for RF-C expression. Although the signal for the His-tagged subunits was clearly present in the protein fraction which was loaded onto the column and corresponded to the size expected for Hisp37 and Hisp40, the complex was not able to bind to the nickel-NTA silica in the spin column. This fact was confirmed with different buffer solutions (Tris- and phosphate-based buffer solutions, see “Materials and Methods”). It also remained unclear whether the bound and eluted 6 x His-tagged subunits represented individual subunits or RF-C subcomplexes and whether their presence was due to incomplete complex formation or complex disintegration. The fact that the RF-C complex did not bind to the nickel-NTA silica in the spin column suggested that either the histidine residues of both tagged subunits were not accessible in the RF-C complex or too short a period of time (centrifugation for 2 min) was allowed for protein binding.

Therefore the protein minipreparation was repeated with nickel-IDA agarose beads (ProBond^{TM} Nickel-Chelating Resin). The bacterial pellet of a 100 ml culture, where RF-C was expressed in BL21(DE3)-Codon Plus RIL cells from the pET21a/p140/p38/Hisp37/Hisp40/p36 construct, was prepared as described in “Materials and Methods” and the fraction containing the soluble cell proteins was incubated with nickel-IDA agarose beads at 4 °C on a roller-shaker for 1 hr. The beads were washed and the bound protein was eluted with 250 mM imidazole in the elution buffers. All protein fractions were again analysed by SDS-PAGE and subsequent western blotting against p140 and the 6 x His-tags. As shown in Figure 13, full-length p140 and 6 x His-tagged subunits of the expected size were detected in elution 1. A decrease of both the p140 and the 6 x His signal was observed in the flow through fraction compared to the load. Additionally, degraded products of p140 did not bind to the beads and were therefore found in the flow through and wash fractions. These results strongly suggested that the five RF-C subunits were indeed able to form a complex, since

p140 was bound to the nickel-IDA resin although it did not carry a 6 x His-tag of its own. The concentration factor for 6 x His-tagged subunits seemed to be manifolds higher in this purification scheme compared to the concentration factor for p140. A possible explanation for this observation might be the fact that individual Hisp37 and Hisp40 subunits or RF-C subcomplexes bound to the nickel-IDA resin with high affinity.

As only the five subunit RF-C complex is enzymatically active, it had to be one of the main goals of recombinant RF-C purification to separate the complex from single subunits and subcomplexes and prevent complex disintegration during the purification procedure by buffer optimisation.

Figure 13 Western blot analysis of a protein minipreparation using nickel-IDA agarose beads



Protein minipreparation with nickel-IDA agarose beads was performed as described above. Two different primary antibodies were used for this western blot analysis: An anti-p140 (upper panel) and an anti-6 x His-antibody (lower panel).

5.8.3 Protein minipreparation: Phosphocellulose

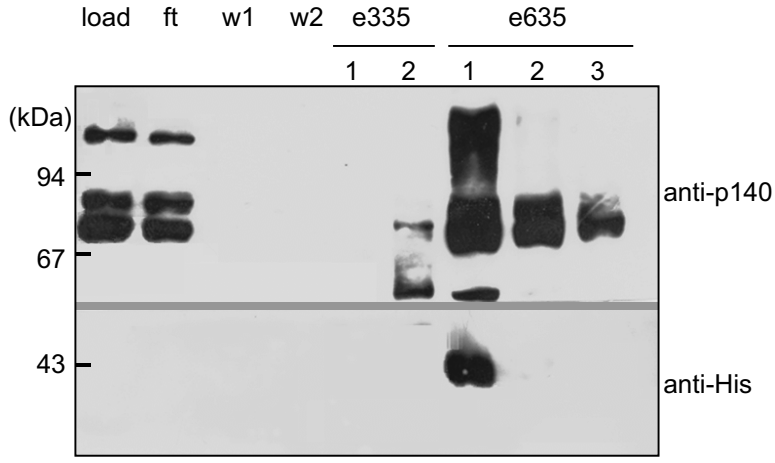
When RF-C was first purified from mammalian cell extracts the first column used in the original purification protocol was a phosphocellulose column (Tsurimoto and Stillman, 1989). DNA binding proteins strongly bind to the phosphate groups in the column material and this bond is disrupted only by high salt concentration. In the case of RF-C the protein complex does not elute at 330 mM salt, but it appears in the 660 mM salt elution. A disadvantage of phosphocellulose is the fact that numerous proteins can bind to DNA and therefore also to phosphocellulose. Since proteins bind to phosphocellulose according to their functional property, this purification step was chosen to complement the nickel-IDA agarose purification step, which bases completely on structural characteristics. It was also interesting to see

whether recombinant RF-C behaved in the same way during the purification procedure with phosphocellulose as RF-C deriving from mammalian cells.

The bacterial pellet of a 100 ml culture, where RF-C was expressed in BL21(DE3)-Codon Plus RIL cells from the pET21a/p140/p38/Hisp37/Hisp40/p36 construct, was prepared as described in “Materials and Methods”. The fraction containing the soluble cell proteins was added to phosphocellulose and the resulting slurry was incubated at 4 °C on a roller-shaker for 2 hrs. The slurry was washed and the bound protein was eluted in two steps with an increasing salt concentration in the elution buffers, first a salt concentration of 335 mM (two elutions), then 635 mM (three elutions). All the protein fractions from this purification procedure were separated by SDS-PAGE and subsequently analysed by western blotting against p140 and the 6 x His-tags.

The result of this experiment is shown in Figure 14. Following the signal for p140 through the purification, it was only a little reduced in the flow through fraction compared to the load, but some RF-C had bound to phosphocellulose and eluted in the 635 mM salt elution, exactly like RF-C in the original purification protocol. The signal for the 6 x His-tagged subunits appeared together with p140 in the 635 mM salt elution fraction, although no 6 x His-signal was detected in the load fraction. This might be due to the fact that only 5 μl of the soluble protein fraction was loaded onto the SDS-polyacrylamide gel because of its high protein concentration in order not to overload the gel.

Figure 14 Western blot analysis of a protein minipreparation using phosphocellulose



Protein minipreparation with phosphocellulose was performed as described above. Two different primary antibodies were used for this western blot analysis: An anti-p140 (upper panel) and an anti-6 x His-antibody (lower panel).

An important disadvantage of this purification step was the fact that it was not possible to separate RF-C from the degraded forms of p140, which had been achieved with nickel-IDA agarose beads. Moreover, as it was already discussed for the nickel purification protocol, RF-C subcomplexes or individual subunits might be present in the bacterial extract alongside the functional five subunit RF-C complex. They were very likely not able to bind to phosphocellulose and were therefore expected to be detected by western blot analysis in the flow through and wash fractions. The fact that the active RF-C complex could not be distinguished from the co-existence of RF-C subcomplexes and individual subunits by western blot analysis turned out to be a major obstacle for the interpretation of the purification data.

5.8.4 Large scale RF-C purification: Optimised protocol for phosphocellulose / nickel combined batch-purification

In the initial protein minipreparation experiments both methods had shown significant disadvantages. First, RF-C had only a weak affinity to nickel-IDA agarose beads although the binding conditions had been optimised. Then, individual 6 x His-tagged RF-C subunits had a strong affinity to nickel and could not be separated from the RF-C complex. On the other hand using phosphocellulose the RF-C complex was very likely separated from subcomplexes and individual subunits, but not from bacterial DNA-binding proteins. Possibly, the *E. coli* homologue of RF-C, the γ-complex, might even co-purify with recombinant human RF-C during the phosphocellulose purification step.

It was therefore obvious to combine the advantages of both purification systems, the strong affinity of the recombinant RF-C complex to phosphocellulose and the high selectivity of the nickel-IDA agarose beads to bind 6 x His-tagged proteins, to an easy and fast, but still efficient purification protocol.

The combined purification is described in detail in “Materials and Methods”. A 1 L culture, where RF-C was expressed in BL21(DE3)-Codon Plus RIL cells from the pET21a/p140/p38/Hisp37/Hisp40/p36 construct according to the optimised expression protocol, was used to prepare the fraction containing the soluble cell proteins. As a first purification step phosphocellulose seemed ideal to pre-clear RF-C from more than 95 % of bacterial proteins. Purification was performed according to the protocol in “Materials and Methods”, the two 635 mM salt elution fractions were pooled and diluted with PC dilution buffer to lower the salt concentration to 300 mM for optimal binding conditions in the next purification step. The nickel purification step was designed with the goal of a high concentration factor for RF-C. The diluted elution fractions were incubated with 200 µl nickel-IDA agarose beads, the beads were washed and the bound protein was eluted in three steps with an increasing concentration of imidazole in the elution buffers. All the fractions were analysed by SDS-PAGE and subsequent Coomassie-blue staining of the SDS-

polyacrylamide gel or western blotting against p140 or the 6 x His-tagged subunits. Additionally, the protein concentration of each fraction was determined to calculate the amount of purified RF-C. A selection of protein fractions was again separated by SDS-PAGE and silver-stained to obtain more precise bands for all five RF-C subunits than it was achieved with Coomassie-blue staining. Finally, the 500 mM elution fraction was analysed by western blotting to prove that the three small subunits Hisp40, Hisp37 and p36 were present in this fraction. For this experiment polyclonal antibodies which had been raised in rabbit (p40 and p37) or chicken (p36) were used to detect these proteins. For p38 detection no antibody was available. The results of all these experiments are shown in Figure 15.

On the Coomassie-blue stained SDS-polyacrylamide gels no bands corresponding to any of the five RF-C subunits could be identified in the load and the phosphocellulose elution fraction, whereas in the nickel elution fraction several bands of the size of the four small subunits and a double band, the upper of the size of p140, were visible (Figure 15 B). The western blot results confirmed that the double band was recognised by the anti-p140 antibody, suggesting that the full-length p140 corresponded to the upper band and the lower appeared to be a degraded form of the large subunit (Figure 15 C). The signal for full-length p140 appeared mainly in the 100 mM and 500 mM imidazole elutions. In the early nickel elutions an even more degraded form of p140 was observed, but like in previous protein minipreparations with nickel-IDA agarose beads it could be separated from full-length p140. Unfortunately, due to the concentration factor of the nickel purification step some degradation was still detectable in the purest RF-C fraction (nickel 500 mM imidazole elution) by western blot analysis and on the Coomassie-blue and silver stained gels. The signals for the two 6 x His-tagged subunits were detected in several fractions. During phosphocellulose purification most of the 6 x His-tagged protein eluted in the first 635 mM salt fraction and from nickel it was primarily eluted with 100 mM and 500 mM imidazole. In the nickel elution fractions the signal was so strong that it was impossible to distinguish between Hisp37 and Hisp40. But in the phosphocellulose wash 1 fraction where individual, surplus subunits were found in a low concentration, a double band was observed. This fact was also confirmed for nickel elution fractions in optimisation experiments using nickel-IDA agarose beads, where the concentration factor for RF-C was considerably lower than in this experiment (data not shown).

The main point of interest in the silver gel was to distinguish the four small RF-C subunits to prove that they were all present in the purified RF-C complex. Three bands of approximately 36 – 43 kDa were clearly visible on the gel (Figure 15 D). Since a 6 x His-tag was added to the p37 and p40 subunits, these two subunits ran higher on the gel compared to the respective endogenous RF-C subunits. It was therefore very likely that Hisp37 ran in exactly the same position on the gel as p38 and only one band for both subunits could be detected.

Finally, the western blot analysis of the 500 mM elution fraction proved the co-existence of the p36, the Hisp37, the Hisp40 and the p140 subunit in the final elution fraction (Figure 15 E). The additional faint bands on the p36 and Hisp37 membranes most likely represented a cross-reaction of the polyclonal antibodies with another small subunit because of the amino-acid sequence similarity of the four small RF-C subunits.

Taken together, these results showed that recombinant RF-C was purified, although some degradation of p140 was still present in the purest fraction. The main obstacle during the purification of recombinant RF-C was the separation of the five subunit RF-C complex from individual subunits, RF-C subcomplexes and p140 degradation products. It was therefore not surprising that the yield of pure RF-C in the form of the five subunit complex was low: Per 1 L of bacterial culture 85 μg of RF-C were purified in the purest fraction (nickel 500mM elution) and an additional 240 μg of relatively pure RF-C were collected in the second and third purest fractions (nickel 100 mM elutions 1 and 2).

Figure 15 Large scale purification of the heteropentameric recombinant human RF-C complex: Phosphocellulose / nickel combined batch-purification

A: Course of the protein concentration in the fractions collected during purification

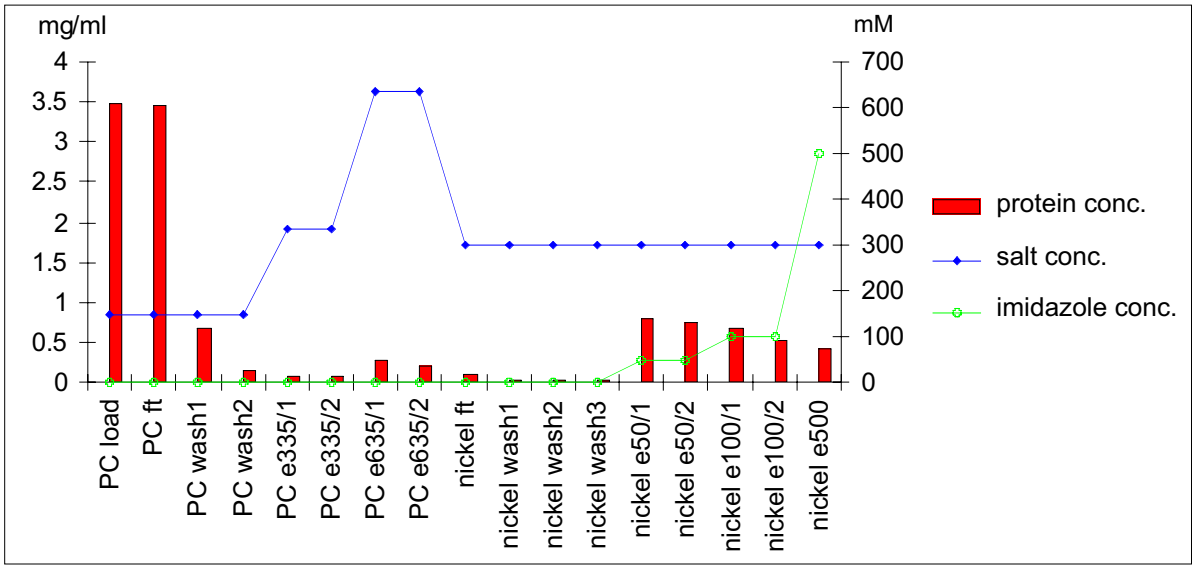
B: SDS-PAGE of the fractions collected during purification (Coomassie-blue stained)

C: Western blot analysis of the fractions collected during purification

D: SDS-PAGE of selected fractions (silver stained)

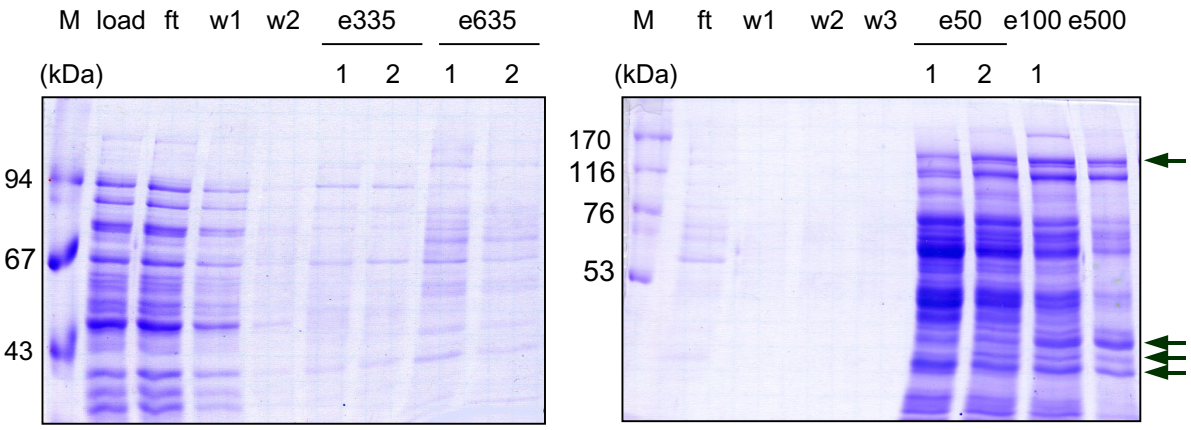
E: Western blot analysis of the 500 mM elution fraction using antibodies against four different RF-C subunits

A



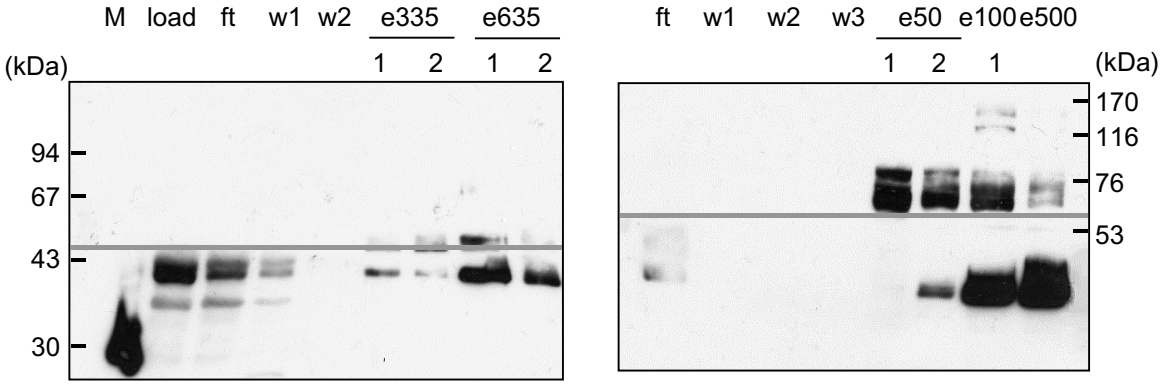
The protein concentration of each fraction was determined by Bradford assay and is here set in relation to the salt or imidazole concentration of the elution buffers.

B

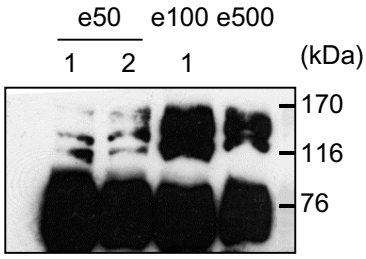


The fractions of the combined phosphocellulose / nickel purification procedure were analysed by SDS-PAGE (10 % gels) and the gels stained with Coomassie-blue. The gel on the left shows the phosphocellulose fractions, the one on the right the nickel fractions.

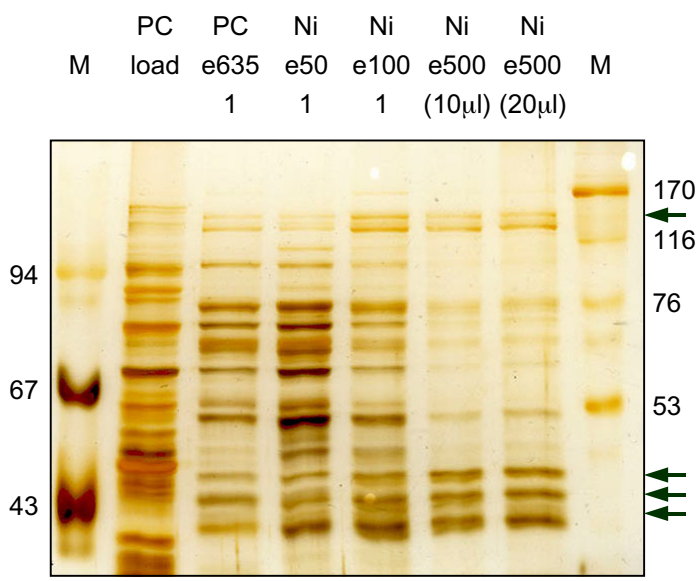
C



Western blot analysis of the fractions which had been collected during the combined phosphocellulose / nickel purification: In the upper panel the membranes were treated with an antibody against p140, in the lower panel with an antibody against 6 x His. Short exposures (10 sec) of the membranes (phosphocellulose fractions on the left, nickel fractions on the right) are shown above. The grey lines indicate where the membrane was cut horizontally. Because the signal for p140 is much weaker than for 6 x His, a longer exposure (20 min) of the upper membrane containing the nickel elution fractions is shown on the right.

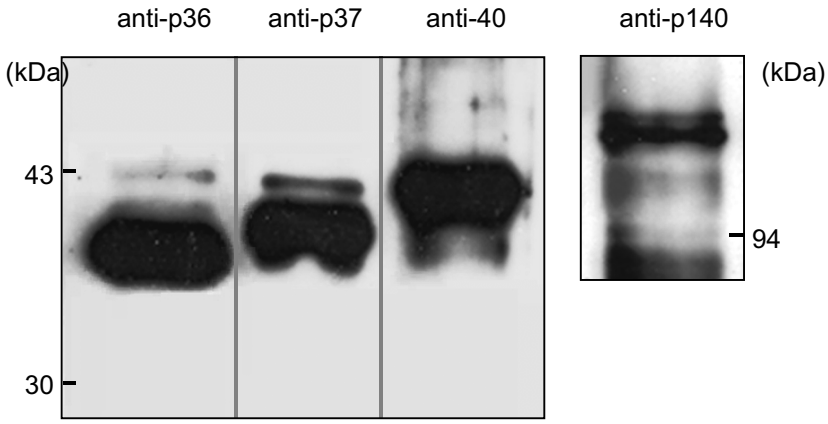


D



Selected fractions of the combined phosphocellulose / nickel purification were loaded onto a 10 % SDS-polyacrylamide gel. The gel was subsequently silver stained as described in “Materials and Methods”. The arrows indicate the positions of p140 and the small RF-C subunits, respectively.

E



Western blot analysis of the 500 mM elution fraction. The following antibodies were used: polyclonal chicken-α p36, polyclonal rabbit-α p37, polyclonal rabbit-α p40 and monoclonal mouse-α p140. The vertical grey lines indicate where the membranes were cut.

5.9 Activity assays

5.9.1 DNA polymerase holoenzyme assay (on a circular template)

During the purification procedure all fractions were tested for specific RF-C activity. The assay described here was quick and specific for RF-C activity and was therefore used to monitor the course of the RF-C activity curve throughout purification. The mechanism of the assay based on the assembly of an active DNA polymerase holoenzyme complex (pol holoenzyme complex) on a primed circular DNA template. The holoenzyme complex consists of the core DNA polymerase (pol δ), the processivity factor PCNA, which has to be loaded onto the circular DNA template by RF-C, and the single-strand DNA-binding protein (SSB). The assay was performed in the presence of ATP, which is required by RF-C to load PCNA onto the DNA and a mix of partially [^{3}H]-labelled deoxyribonucleotides (dATP, dGTP, dCTP, dTTP and [^{3}H]dTTP). The reaction mix was incubated at 37 °C for 30 min to allow DNA synthesis by the holoenzyme complex. Afterwards, incorporated [^{3}H]-labelled nucleotides were measured as described in “Materials and Methods”. To distinguish between radioactivity which was incorporated by the RF-C-dependent holoenzyme complex and e.g. bacterial polymerases which are capable of RF-C independent DNA synthesis the assay was performed in two different compositions for each purification fraction. One contained all the factors of the holoenzyme complex, the other assay mix was incubated lacking pol δ. The results for selected fractions of the optimised phosphocellulose / nickel purification are shown in Figure 16 B.

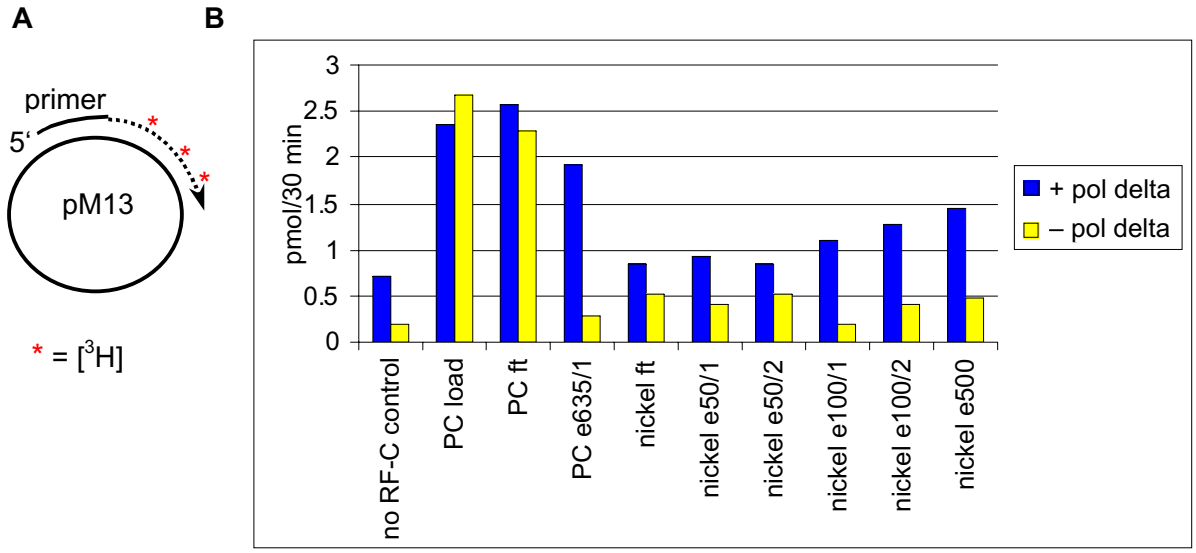
Additionally, the fractions were diluted several folds in nickel elution buffer containing no imidazole and the different dilutions were tested. The results of two of the fractions, the 100 mM imidazole elution 1, which contained RF-C but also a considerable amount of protein contamination, and the 500 mM imidazole elution, which was the purest RF-C fraction, are shown in Figure 16 C. For the 100 mM fraction an increase of RF-C-activity was observed, which was followed by a sharp decrease at higher concentrations. This might suggest that at higher concentrations the negative effect of a contamination became predominant. On the other hand, in the 500 mM elution a concentration-dependent increase of activity was seen at higher amounts of RF-C.

Figure 16 DNA polymerase holoenzyme assay (on a circular template)

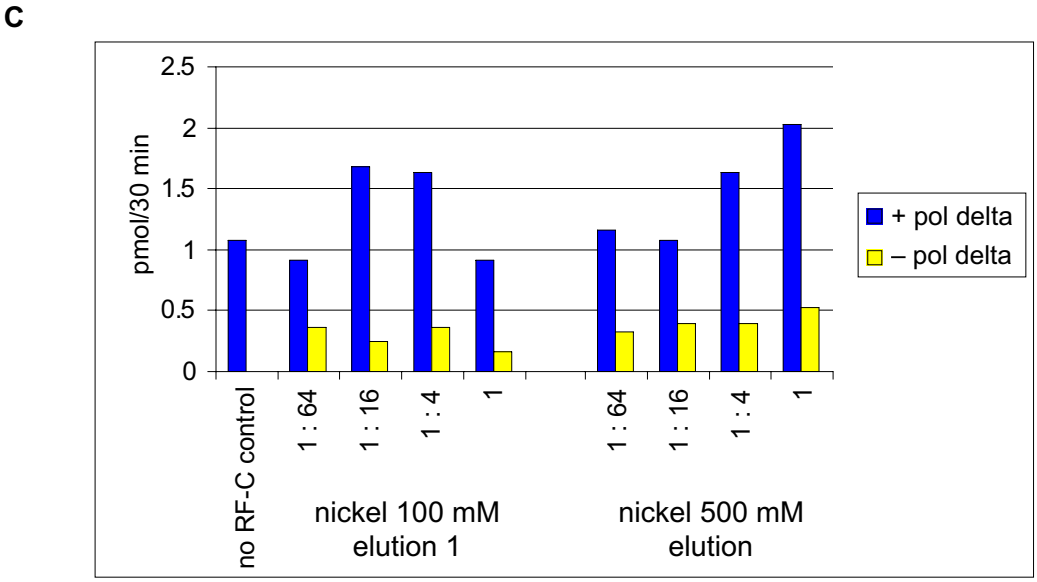
A: DNA template used in this assay

B: Assay for selected fractions of the combined phosphocellulose / nickel purification

C: Assay showing the concentration-dependent stimulation of DNA synthesis by recombinant human RF-C



DNA polymerase holoenzyme assay for selected fractions of the combined phosphocellulose / nickel purification. Each fraction was tested in two separate assays, one with and the other without pol δ to discriminate between RF-C activity and the polymerase activity of bacterial DNA polymerases. For each assay 1 µl of the corresponding RF-C fraction was used.



DNA polymerase holoenzyme assay: The two RF-C elution fractions showing the highest RF-C activity in Figure 16 B were diluted in nickel elution buffer without imidazole in 1 : 4 steps. 1 µl of each dilution was added to the assay.

5.9.2 DNA polymerase holoenzyme assay (on a linear template)

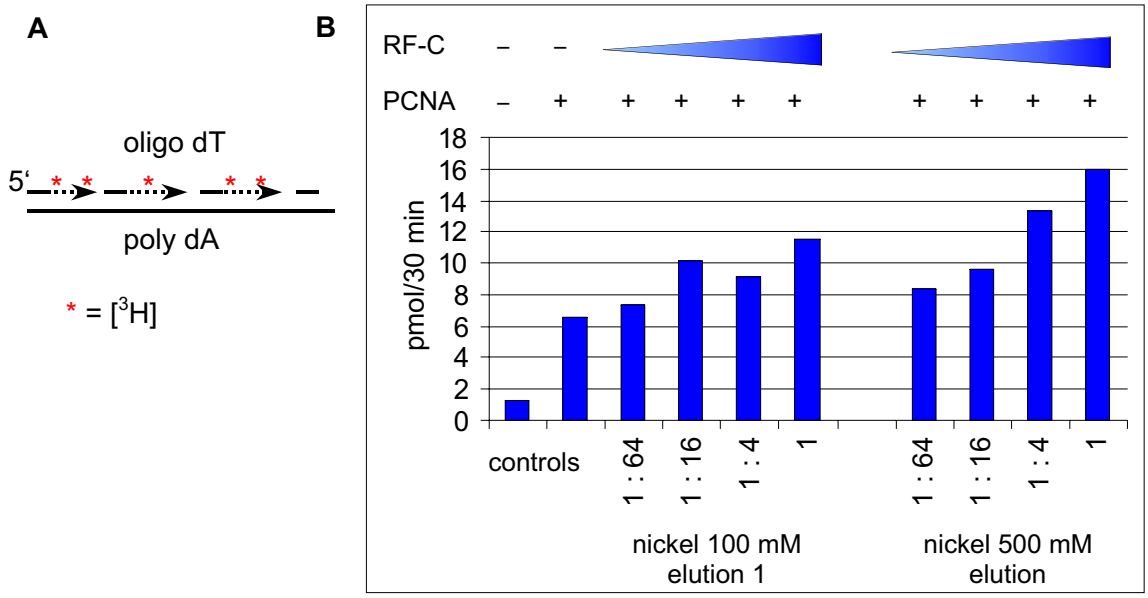
This assay was performed in a modified form of the common polymerase activity assay on a poly dA / oligo dT DNA template. Since a linear DNA template was used in these assays, PCNA-dependent polymerase activity was measured in the absence of RF-C because PCNA could easily slide onto and off the DNA template, but only when the pH of the reaction buffer was lowered to 6.5. In the modified assay for RF-C activity a pH of 7.5 in the reaction buffer prevented the assembly of PCNA on the negatively charged DNA template, because the inside of the protein ring then carried an increased negative charge. DNA synthesis could therefore be stimulated by RF-C-dependent loading of PCNA at primer-template junctions. It has to be emphasised that this assay showed a relatively high background activity, due to PCNA sliding onto the DNA template, and RF-C activity had to be measured in comparison to the assay control lacking RF-C.

The assay was performed in parallel to the DNA polymerase holoenzyme assay shown in Figure 16 C. The same fractions and dilutions were tested and the assay was performed as described in “Materials and Methods”. The results of this experiment are shown in Figure 17. The undiluted RF-C 100 mM elution stimulated DNA synthesis 1.75 fold, whereas the RF-C 500 mM elution achieved a stimulation of 2.4 and was again concentration-dependent.

Figure 17 DNA polymerase holoenzyme assay (on a linear template)

A: DNA template used in this assay

B: Assay showing the concentration-dependent stimulation of DNA synthesis by recombinant human RF-C



Modified DNA polymerase activity assay on a poly dA / oligo dT template to measure the stimulatory effect of RF-C on PCNA-dependent DNA synthesis by pol δ.

5.9.3 PCNA loading assay

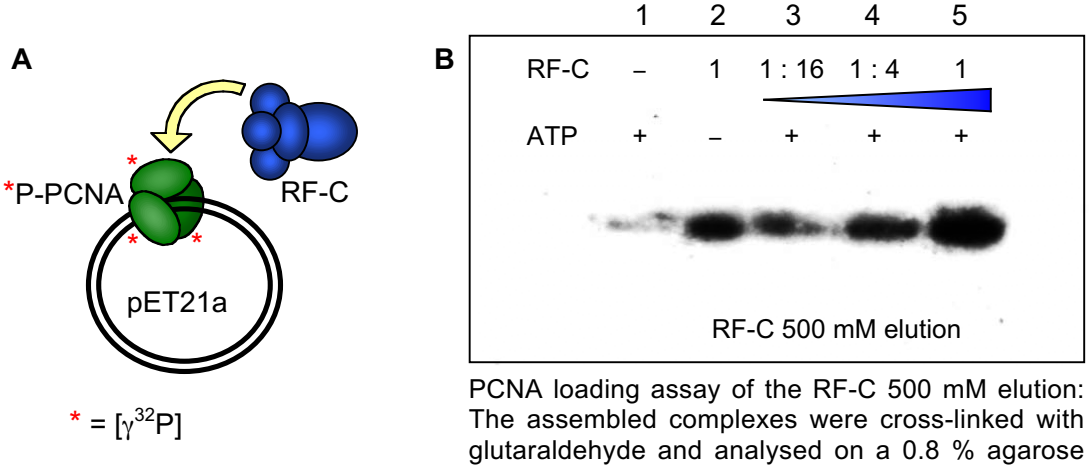
The properties of recombinant RF-C were further characterised by a PCNA loading assay, by which the ATP-dependent clamp loading activity of RF-C could be assessed. Furthermore this assay did not base on the indirect measuring of incorporated labelled nucleotides, but directly visualised the labelled PCNA clamps which had been loaded onto a DNA template.

In a first step phosphorylatable PCNA (ph-PCNA) was phosphorylated by the bovine heart muscle kinase using [γ^{32}P]-ATP. Then, 1 µl of the nickel 500 mM elution or a dilution of the fraction was incubated with 100 ng of circular double-stranded DNA (pET21a plasmid) and 50 ng of phosphorylated PCNA. The assay was performed in the presence of 1 mM ATP and, to prove the ATP-dependency of the clamp loading process, the undiluted elution fraction was additionally tested in a reaction without ATP. After 3 min of incubation at 37 °C the assembled complexes were cross-linked with glutaraldehyde and then immediately loaded onto a 0.8 % agarose gel. The autoradiography of the dried gel is shown in Figure 18. A concentration-dependent increase of the signal was observed (lanes 3 – 5). The reaction was dependent on RF-C (lane 1). The only slight dependency on added ATP was due to endogenous ATP deriving from the PCNA labelling reaction and the RF-C preparation. During purification ATP had to be included in the wash and elution buffers in order to obtain an active recombinant RF-C complex (U. Hübscher, personal communication).

Figure 18 PCNA loading assay

A: DNA template and *P-PCNA used in this assay

B: PCNA loading assay: Concentration and ATP-dependent PCNA loading by recombinant human RF-C



PCNA loading assay of the RF-C 500 mM elution: The assembled complexes were cross-linked with glutaraldehyde and analysed on a 0.8 % agarose gel. The autoradiography of the dried gel showed that the clamp loading process was RF-C concentration- and ATP-dependent.

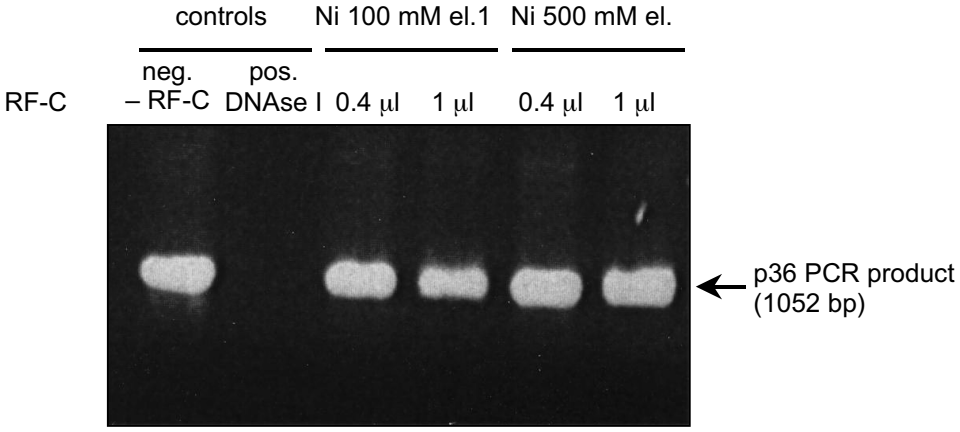
In summary, both the two DNA polymerase holoenzyme assays (on a circular and on a linear template) and the PCNA loading assay suggested that the recombinant human five subunit RF-C complex was enzymatically active.

5.9.4 Exo- / Endonuclease assay

To rule out the possibility that the DNA template used in the RF-C activity assays was digested by bacterial exo- or endonucleases, the nickel 100 mM elution 1 and the nickel 500 mM elution were tested for this kind of contamination. This was absolutely necessary, because recombinant proteins purified from bacteria are more likely contaminated by nucleases than from any other source. Recombinant RF-C will mainly be used for *in vitro* clamp loading experiments, which of course contain DNA. Even the slightest nuclease contamination would make recombinant RF-C useless for any of these experiments.

To detect nuclease contaminants 0.4 µl and 1 µl of the nickel 100 mM elution 1 and the 500 mM elution were incubated with 200 ng of linear, double stranded DNA at 37 °C for 1hr. The PCR product of the p36 gene amplification (1052 bp) was used for this purpose. The assay was performed as described in “Materials and Methods”. As a nuclease positive control 200 u of DNAse I were incubated with the same amount of DNA as mentioned above. The negative control sample contained neither RF-C nor DNAse I and was also incubated at 37 °C for 1 hr. After incubation the samples were immediately loaded onto a 0.8 % agarose gel containing ethidiumbromide. The results of this experiment are shown in Figure 19. When DNAse I was added to the reaction mix, the DNA was completely digested resulting in the absence of a DNA signal on the agarose gel. On the contrary, when the RF-C fractions were added no reduction of the DNA signal was observed. These results suggested that both the nickel 100 mM elution 1 and the nickel 500 mM elution fraction were completely free from bacterial exo- and endonucleases.

Figure 19 Exo- / Endonuclease assay



Exo- / Endonuclease assay for the two RF-C fractions Nickel 100 mM elution 1 and Nickel 500 mM elution. 0.4 µl or 1 µl of the fractions or 200 u of DNAse I were incubated with 200 ng of linear, double stranded DNA at 37 °C for 1 hr. The samples were analysed on a 0.8 % agarose gel containing ethidiumbromide.

6 Discussion

Ever since the discovery of RF-C in 1989, many groups had tried to establish an efficient system for the production of recombinant RF-C. Although RF-C was expressed in most of the systems, all these attempts proved to be either not practical or too expensive for the production of large amounts of recombinant RF-C. The easiest and cheapest system remains the production of heterologous proteins by an *E. coli* host strain. Despite its obvious advantages the bacterial system also has disadvantages, which can lead to impassable obstacles in the overexpression of difficult target proteins. RF-C is known to be an extremely difficult target protein, because individually expressed subunits are completely insoluble (Chen, et al., 1992a, Chen, et al., 1992b, Luckow, et al., 1994). Moreover, when the five yeast RF-C subunits were co-expressed in *E. coli*, no polypeptide corresponding to the large subunit was detected even after optimisation of the induction conditions. Only when a N-terminally truncated form of the large subunit was co-expressed with the four small subunits, a modified RF-C complex was formed and could be purified (Gomes, et al., 2000). Neither yeast nor human RF-C has ever been overexpressed in *E. coli* in its full-length form.

It was therefore the goal of this thesis project to establish and optimise a bacterial overexpression system for full-length human RF-C, taking advantage of the recently developed *E. coli* host strains which had been improved for difficult target proteins. These *E. coli* BL21(DE3)-Codon Plus RIL and RP strains overexpress tRNAs which are rarely used in *E. coli*, the argU, ileY and the leuW tRNAs in the RIL strain and the argU and the proL tRNAs in the RP strain. These extra tRNAs prevent the issue of codon-bias, which can inhibit the synthesis of recombinant proteins or lead to the production of inactive enzymes because of the misincorporation of amino-acids.

The five RF-C genes were therefore all cloned into a pET21a vector under the control of the strong T7 promoter. For efficient and equal transcription and translation of all RF-C subunits their genes were aligned on the vector in a way that a long mRNA was produced with the information for all five subunits. Additionally, the gene for each RF-C subunit carried its own bacterial ribosome binding site upstream of the gene. This plasmid construct was expressed in an *E. coli* BL21(DE3)-Codon Plus RIL strain. It became soon clear that only under thoroughly optimised induction conditions a sufficient level of expression could be achieved for RF-C. Lowering the induction temperature down to 23 °C solved the two major problems of induction: The expression level of soluble RF-C was dramatically increased, whereas degradation of the large RF-C subunit (p140) was considerably reduced. The following induction conditions were found to yield a maximal amount of soluble RF-C: The bacterial culture was grown to a high optical density (OD_{600} = 1) to minimise the risk of RF-C interfering with the

bacterial life cycle. Consequently, a high concentration of IPTG (1 mM) was used to induce RF-C expression. Finally an induction period of 3 hrs at 23 °C was found to be optimal to reach the maximal level of expression with an acceptable level of p140 degradation.

Next, a purification scheme was developed and optimised for the purification of recombinant RF-C. Two of the small RF-C subunits carried an N-terminal 6 x His-tag (Hisp37 and Hisp40). Hence the binding affinity of recombinant RF-C to nickel-IDA agarose beads was tested. Additionally, as phosphocellulose was the first purification step of the original RF-C purification protocol (Tsurimoto and Stillman, 1989), it was tested whether recombinant RF-C also had the ability of binding to phosphocellulose. Neither purification scheme yielded RF-C fractions which were pure enough for the use in enzymatic assays. Therefore the two schemes were linked in order to combine the advantages of both systems. In a first step the bacterial extract was pre-cleared with phosphocellulose making use of the functional property of RF-C to bind to DNA. More than 95 % of *E. coli* proteins and surplus individual RF-C subunits were removed with this step. Then, the structural property of recombinant RF-C to carry two 6 x His-tags was utilised for the purification step with nickel-IDA agarose beads. The combination of two purification schemes which were able to separate proteins according to two different properties were found to be extremely efficient for the purification of recombinant RF-C from *E. coli* cell extract.

The main problem during purification was the existence of RF-C subcomplexes, individual subunits and degradation products, especially of the p140 subunit, alongside the active five subunit RF-C complex and their separation. Especially during nickel purification disintegration of the five subunit RF-C was often observed in form of the loss of the large subunit, which then appeared in earlier fractions than the 6 x His-tagged subunits. This observation might reflect the findings of recent studies that under certain circumstances the p140 subunit of the RF-C complex can be replaced by other proteins, e. g. Rad17 (Griffith, et al., 2002, Lindsey-Boltz, et al., 2001, Shiomi, et al., 2002, Venclovas, et al., 2002). It was also speculated that the four small subunits of RF-C form a tight complex whereas the p140 subunit is only loosely attached.

The analysis of the purest RF-C fraction by western blot proved that the p140, the Hisp40, the Hisp37 and the p36 subunits were present in this fraction. For p38 detection no antibody was available. This result strongly indicated that the five subunit RF-C complex was formed and purified successfully, because p36 and p140 bound to nickel-ions without being 6 x His-tagged themselves. Although the elution fraction was not tested for the presence of p38, this subunit had to be present because it had been shown previously that p140 can only form a complex with the small RF-C subunits in the presence of the linker subunit p38 (Cai, et al., 1997, Podust and Fanning, 1997, Uhlmann, et al., 1996, Uhlmann, et al., 1997b). After all,

recombinant RF-C did not abandon the properties of endogenous RF-C which make it an extremely delicate and difficult enzyme. It was therefore expected that the yield of pure RF-C in the form of the five subunit complex was quite low: Per 1 L of bacterial culture 85 μg of RF-C were purified in the purest fraction (nickel 500 mM elution) and an additional 240 μg of relatively pure RF-C were collected in the second and third purest fractions (nickel 100 mM elutions 1 and 2).

Finally, the data suggested that recombinant RF-C had similar enzymatic properties as endogenous RF-C (Fairman, et al., 1989, Podust, et al., 1995). The enzymatic activities of recombinant RF-C were analysed with several assays. The stimulatory effect of recombinant RF-C on DNA synthesis by pol δ was measured with a DNA polymerase holoenzyme activity assay on either a circular or a linear DNA template. Then the ATP-dependent clamp loading activity of recombinant RF-C was confirmed with a PCNA loading assay, where [γ^{32}P]-labelled PCNA was loaded onto a circular, double-stranded DNA.
It was also shown that the fractions containing pure RF-C were free of bacterial exo- and endonucleases, which is an absolute necessity for the future application of recombinant RF-C in *in vitro* replication or repair activity assays which contain DNA.

In this thesis project a completely new approach to produce human RF-C in a recombinant system was established and optimised successfully. This cheap and quick expression and purification system has not only made large quantities of enzymatically active RF-C available. The progresses made in this project to simultaneously overexpress multi-subunit enzymes in special *E. coli* expression hosts can also serve as a basis for the bacterial overexpression of other difficult multi-subunit complexes like the novel RF-C complex (Rad17 and the four small RF-C subunits) or the Rad9-Rad1-Hus1 (9-1-1) complex.

7 References

Allen, B. L., Uhlmann, F., Gaur, L. K., Mulder, B. A., Posey, K. L., Jones, L. B. and Hardin, S. H. (1998) DNA recognition properties of the N-terminal DNA binding domain within the large subunit of replication factor C. *Nucleic Acids Res*, 26, 3877-82

Anderson, L. A. and Perkins, N. D. (2002) The large subunit of replication factor C interacts with the histone deacetylase, HDAC1. *J Biol Chem*, 277, 29550-4

Anderson, L. A. and Perkins, N. D. (2003) Regulation of RelA (p65) Function by the Large Subunit of Replication Factor C. *Mol Cell Biol*, 23, 721-32

Bork, P., Hofmann, K., Bucher, P., Neuwald, A. F., Altschul, S. F. and Koonin, E. V. (1997) A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. *FASEB J*, 11, 68-76

Burbelo, P. D., Utani, A., Pan, Z. Q. and Yamada, Y. (1993) Cloning of the large subunit of activator 1 (replication factor C) reveals homology with bacterial DNA ligases. *Proc Natl Acad Sci U S A*, 90, 11543-7

Burgers, P. M. (1991) *Saccharomyces cerevisiae* replication factor C. II. Formation and activity of complexes with the proliferating cell nuclear antigen and with DNA polymerases delta and epsilon. *J Biol Chem*, 266, 22698-706

Burgers, P. M. (1999) Overexpression of multisubunit replication factors in yeast. *Methods*, 18, 349-55

Cai, J., Gibbs, E., Uhlmann, F., Phillips, B., Yao, N., O'Donnell, M. and Hurwitz, J. (1997) A complex consisting of human replication factor C p40, p37, and p36 subunits is a DNA-dependent ATPase and an intermediate in the assembly of the holoenzyme. *J Biol Chem*, 272, 18974-81

Cai, J., Uhlmann, F., Gibbs, E., Flores-Rozas, H., Lee, C. G., Phillips, B., Finkelstein, J., Yao, N., O'Donnell, M. and Hurwitz, J. (1996) Reconstitution of human replication factor C from its five subunits in baculovirus-infected insect cells. *Proc Natl Acad Sci U S A*, 93, 12896-901

Cai, J., Yao, N., Gibbs, E., Finkelstein, J., Phillips, B., O'Donnell, M. and Hurwitz, J. (1998) ATP hydrolysis catalyzed by human replication factor C requires participation of multiple subunits. *Proc Natl Acad Sci U S A*, 95, 11607-12

Canman, C. E. (2001) Replication checkpoint: preventing mitotic catastrophe. *Curr Biol*, 11, R121-4

Caruthers, J. M. and McKay, D.B. (2002) Helicase structure and mechanism. *Curr Opin Struct Biol*, 12, 123-33

Chen, M., Pan, Z. Q. and Hurwitz, J. (1992a) Sequence and expression in *Escherichia coli* of the 40-kDa subunit of activator 1 (replication factor C) of HeLa cells. *Proc Natl Acad Sci U S A*, 89, 2516-20

Chen, M., Pan, Z. Q. and Hurwitz, J. (1992b) Studies of the cloned 37-kDa subunit of activator 1 (replication factor C) of HeLa cells. *Proc Natl Acad Sci U S A*, 89, 5211-5

Cullmann, G., Fien, K., Kobayashi, R. and Stillman, B. (1995) Characterization of the five replication factor C genes of *Saccharomyces cerevisiae*. *Mol Cell Biol*, 15, 4661-71

Ellison, V. and Stillman, B. (1998) Reconstitution of recombinant human replication factor C (RFC) and identification of an RFC subcomplex possessing DNA-dependent ATPase activity. *J Biol Chem*, 273, 5979-87

Fairman, M. P., Prelich, G., Tsurimoto, T. and Stillman, B. (1989) Replication of SV40 in vitro using proteins derived from a human cell extract. *J Cell Sci Suppl*, 12, 161-9

Fien, K. and Stillman, B. (1992) Identification of replication factor C from *Saccharomyces cerevisiae*: a component of the leading-strand DNA replication complex. *Mol Cell Biol*, 12, 155-63

Fotedar, R. and Fotedar, A. (1995) Cell cycle control of DNA replication. *Prog Cell Cycle Res*, 1, 73-89

Fotedar, R., Mossi, R., Fitzgerald, P., Rousselle, T., Maga, G., Brickner, H., Messier, H., Kasibhatla, S., Hubscher, U. and Fotedar, A. (1996) A conserved domain of the large subunit of replication factor C binds PCNA and acts like a dominant negative inhibitor of DNA replication in mammalian cells. *EMBO J*, 15, 4423-33

Frouin, I., Montecucco, A., Biamonti, G., Hubscher, U., Spadari, S. and Maga, G. (2002) Cell cycle-dependent dynamic association of cyclin/Cdk complexes with human DNA replication proteins. *EMBO J*, 21, 2485-95

Fujise, K., Zhang, D., Liu, J. and Yeh, E. T. (2000) Regulation of apoptosis and cell cycle progression by MCL1. Differential role of proliferating cell nuclear antigen. *J Biol Chem*, 275, 39458-65

Fukuda, K., Morioka, H., Imajou, S., Ikeda, S., Ohtsuka, E. and Tsurimoto, T. (1995) Structure-function relationship of the eukaryotic DNA replication factor, proliferating cell nuclear antigen. *J Biol Chem*, 270, 22527-34

Gerik, K. J., Gary, S. L. and Burgers, P. M. (1997) Overproduction and affinity purification of *Saccharomyces cerevisiae* replication factor C. *J Biol Chem*, 272, 1256-62

Gomes, X. V. and Burgers, P. M. (2001) ATP utilization by yeast replication factor C. I. ATP-mediated interaction with DNA and with proliferating cell nuclear antigen. *J Biol Chem*, 276, 34768-75

Gomes, X. V., Gary, S. L. and Burgers, P. M. (2000) Overproduction in *Escherichia coli* and characterization of yeast replication factor C lacking the ligase homology domain. *J Biol Chem*, 275, 14541-9

Gomes, X. V., Schmidt, S. L. and Burgers, P. M. (2001) ATP utilization by yeast replication factor C. II. Multiple stepwise ATP binding events are required to load proliferating cell nuclear antigen onto primed DNA. *J Biol Chem*, 276, 34776-83

Griffith, J. D., Lindsey-Boltz, L. A. and Sancar, A. (2002) Structures of the human Rad17-replication factor C and checkpoint Rad 9-1-1 complexes visualized by glycerol spray/low voltage microscopy. *J Biol Chem*, 277, 15233-6

Henneke, G., Friedrich-Heinecken, E. and Hubscher, U. (2003a) Flap endonuclease 1: a novel tumour suppresser protein. *Trends Biochem Sci*, 28, 384-90

Henneke, G., Koundrioukoff, S. and Hubscher, U. (2003b) Multiple roles for kinases in DNA replication. *EMBO Report*, 4, 252-6

Hingorani, M. M. and Coman, M. M. (2002) On the Specificity of Interaction between the *Saccharomyces cerevisiae* Clamp Loader Replication Factor C and Primed DNA Templates during DNA Replication. *J Biol Chem*, 277, 47213-24

Holmes, A. M. and Haber, J. E. (1999) Double-strand break repair in yeast requires both leading and lagging strand DNA polymerases. *Cell*, 96, 415-24

Hong, S. H., Park, S. J., Kong, H. J., Shuman, J. D. and Cheong, J. H. (2001) Functional interaction of bZIP proteins and the large subunit of replication factor C in liver and adipose cells. *J Biol Chem*, 276, 28098-28105

Hubscher, U., Maga, G. and Spadari, S. (2002) Eukaryotic DNA polymerases. *Annu Rev Biochem*, 71, 133-63

Jeruzalmi, D., O'Donnell, M. and Kuriyan, J. (2001a) Crystal structure of the processivity clamp loader gamma (gamma) complex of *E. coli* DNA polymerase III. *Cell*, 106, 429-41

Jeruzalmi, D., O'Donnell, M. and Kuriyan, J. (2002) Clamp loaders and sliding clamps. *Curr Opin Struct Biol*, 12, 217-24

Jeruzalmi, D., Yurieva, O., Zhao, Y., Young, M., Stewart, J., Hingorani, M., O'Donnell, M. and Kuriyan, J. (2001b) Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III. *Cell*, 106, 417-28

Jonnson, Z. O., Hindges, R. and Hubscher, U. (1998) Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. *EMBO J*, 17, 2412-25

Jonnson, Z. O. and Hubscher, U. (1997) Proliferating cell nuclear antigen: more than a clamp for DNA polymerases. *Bioessays*, 19, 967-75

Keller, R. C., Mossi, R., Maga, G., Wellinger, R. E., Hubscher, U. and Sogo, J. M. (1999) Electron microscopic analysis reveals that replication factor C is sequestered by single-stranded DNA. *Nucleic Acids Res*, 27, 3433-7

Kelman, Z. (1997) PCNA: structure, functions and interactions. *Oncogene*, 14, 629-40

Kelman, Z. and Hurwitz, J. (2000) A unique organization of the protein subunits of the DNA polymerase clamp loader in the archaeon *Methanobacterium thermoautotrophicum deltaH*. *J Biol Chem*, 275, 7327-36

Kolodner, R. D. and Marsischky, G. T. (1999) Eukaryotic DNA mismatch repair. *Curr Opin Genet Dev*, 9, 89-96

Kolpashchikov, D. M., Hughes, P., Favre, A., Baldacci, G. and Lavrik, O. I. (2001) Localization of the large subunit of replication factor C near the 5'-end of DNA primers. *J Mol Recognit*, 14, 239-44

Krishna, T. S., X.P., K., Gary, S., Burgers, P. M. and Kuriyan, J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell*, 79, 1233-43

Kuriyan, J. and O'Donnell, M. (1993) Sliding clamps of DNA polymerases. *J Mol Biol*, 234, 915-25

Leu, F. P. and O'Donnell, M. (2001) Interplay of clamp loader subunits in opening the beta sliding clamp of *Escherichia coli* DNA polymerase III holoenzyme. *J Biol Chem*, 276, 47185-94

Lindsey-Boltz, L. A., Bermudez, V. P., Hurwitz, J. and Sancar, A. (2001) Purification and characterization of human DNA damage checkpoint Rad complexes. *Proc Natl Acad Sci U S A.*, 98, 11236-41

Luckow, B., Bunz, F., Stillman, B., Lichter, P. and Schutz, G. (1994) Cloning, expression, and chromosomal localization of the 140-kilodalton subunit of replication factor C from mice and humans. *Mol Cell Biol*, 14, 1626-34

Maga, G. and Hubscher, U. (2003) Proliferating cell nuclear antigen (PCNA): a dancer with many partners. *J Cell Sci*, 116, 3051-60

Maga, G., Mossi, R., Fischer, R., Berchtold, M. W. and Hubscher, U. (1997) Phosphorylation of the PCNA binding domain of the large subunit of replication factor C by Ca2+/calmodulin-dependent protein kinase II inhibits DNA synthesis. *Biochemistry*, 36, 5300-10

Maga, G., Stucki, M., Spardari, S. and Hubscher, U. (2000) DNA polymerase switching: I. Replication factor C displaces DNA polymerase alpha prior to PCNA loading. *J Mol Biol*, 295, 791-801

Matsumoto, Y. (2001) Molecular mechanism of PCNA-dependent base excision repair. *Prog Nucleic Acid Res Mol Biol*, 68, 129-38

Melo, J. and Toczyski, D. (2002) A unified view of the DNA-damage checkpoint. *Curr Opin Cell Biol*, 14, 237-45

Mossi, R. and Hubscher, U. (1998) Clamping down on clamps and clamp loaders--the eukaryotic replication factor C. *Eur J Biochem*, 254, 209-16

Mossi, R., Jonsson, Z. O., Allen, B. L., Hardin, S. H. and Hubscher, U. (1997) Replication factor C interacts with the C-terminal side of proliferating cell nuclear antigen. *J Biol Chem*, 272, 1769-76

Mossi, R., Keller, R. C., Ferrari, E. and Hubscher, U. (2000) DNA polymerase switching: II. Replication factor C abrogates primer synthesis by DNA polymerase alpha at a critical length. *J Mol Biol*, 295, 803-14

Neuwald, A. F., Aravind, L., Spouge, J. L. and Koonin, E. V. (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res*, 9, 27-43

Nigg, E. A. (1995) Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle, *Bioessays*, 17, 471-480

O'Donnell, M., Jeruzalmi, D. and Kuriyan, J. (2001) Clamp loader structure predicts the architecture of DNA polymerase III holoenzyme and RFC. *Curr Biol*, 11, R935-46

Ohta, S., Shiomi, Y., Sugimoto, K., Obuse, C. and Tsurimoto, T. (2002) A proteomics approach to identify proliferating cell nuclear antigen (PCNA)-binding proteins in human cell lysates. Identification of the human CHL12/RFCs2-5 complex as a novel PCNA-binding protein. *J Biol Chem*, 277, 40362-7

Oku, T., Ikeda, S., Sasaki, H., Fukuda, K., Morioka, H., Ohtsuka, E., Yoshikawa, H. and Tsurimoto, T. (1998) Functional sites of human PCNA which interact with p21 (Cip1/Waf1), DNA polymerase delta and replication factor C. *Genes Cells*, 3, 357-69

Pan, Z. Q., Chen, M. and Hurwitz, J. (1993) The subunits of activator 1 (replication factor C) carry out multiple functions essential for proliferating-cell nuclear antigen-dependent DNA synthesis. *Proc Natl Acad Sci U S A*, 90, 6-10

Pascucci, B., Maga, G., Hubscher, U., Bjoras, M., Seeberg, E., Hickson, I. D., Villani, G., Giordano, C., Cellai, L. and Dogliotti, E. (2002) Reconstitution of the base excision repair pathway for 7,8-dihydro-8-oxoguanine with purified human proteins. *Nucleic Acids Res*, 30, 2124-30

Pennaneach, V., Salles-Passador, I., Munshi, A., Brickner, H., Regazzoni, K., Dick, F., Dyson, N., Chen, T. T., Wang, J. Y. J., Fotedar, R. and Fotedar, A. (2001) The large subunit of replication factor C promotes cell survival after DNA damage in an LxCxE motif- and Rb-dependent manner. *Mol. Cell*, 7, 715-727

Pisani, F. M., De Felice, M., Carpentieri, F. and Rossi, M. (2000) Biochemical characterization of a clamp-loader complex homologous to eukaryotic replication factor C from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *J Mol Biol*, 301, 61-73

Podust, L. M., Podust, V. N., Sogo, J. M. and Hubscher, U. (1995) Mammalian DNA polymerase auxiliary proteins: analysis of replication factor C-catalyzed proliferating cell nuclear antigen loading onto circular double-stranded DNA. *Mol Cell Biol*, 15, 3072-8

Podust, V., Mikhailov, V., Georgaki, A. and Hubscher, U. (1992) DNA polymerase delta and epsilon holoenzymes from calf thymus. *Chromosoma*, 102, S133-41

Podust, V. N., Chang, L. S., Ott, R., Dianov, G. L. and Fanning, E. (2002) Reconstitution of human DNA polymerase delta using recombinant baculoviruses: the p12 subunit potentiates DNA polymerizing activity of the four-subunit enzyme. *J Biol Chem*, 277, 3894-901

Podust, V. N. and Fanning, E. (1997) Assembly of functional replication factor C expressed using recombinant baculoviruses. *J Biol Chem*, 272, 6303-10

Podust, V. N., Tiwari, N., Ott, R. and Fanning, E. (1998) Functional interactions among the subunits of replication factor C potentiate and modulate its ATPase activity. *J Biol Chem*, 273, 12935-42

Rauen, M., Burtelow, M. A., Dufault, V. M. and Karnitz, L. M. (2000) The human checkpoint protein hRad17 interacts with the PCNA-like proteins hRad1, hHus1, and hRad9. *J Biol Chem*, 275, 29767-71

Reynolds, N., Fantes, P. A. and MacNeill, S. A. (1999) A key role for replication factor C in DNA replication checkpoint function in fission yeast. *Nucleic Acids Res*, 27, 462-9

Rheaume, E., Cohen, L. Y., Uhlmann, F., Lazure, C., Alam, A., Hurwitz, J., Sekaly, R. P. and Denis, F. (1997) The large subunit of replication factor C is a substrate for caspase-3 in vitro and is cleaved by a caspase-3-like protease during Fas-mediated apoptosis. *EMBO J*, 16, 6346-54

Roos-Mattjus, P., Vroman, B. T., Burtelow, M. A., Rauen, M., Eapen, A. K. and Karnitz, L. M. (2002) Genotoxin-induced Rad9-Hus1-Rad1 (9-1-1) chromatin association is an early checkpoint signaling event. *J Biol Chem*, 277, 43809-12

Schmidt, S. L., Gomes, X. V. and Burgers, P. M. (2001a) ATP utilization by yeast replication factor C. III. The ATP-binding domains of Rfc2, Rfc3, and Rfc4 are essential for DNA recognition and clamp loading. *J Biol Chem*, 276, 34784-91

Schmidt, S. L., Pautz, A. L. and Burgers, P. M. (2001b) ATP utilization by yeast replication factor C. IV. RFC ATP-binding mutants show defects in DNA replication, DNA repair, and checkpoint regulation. *J Biol Chem*, 276, 34792-800

Schurtenberger, P., Egelhaaf, S. U., Hindges, R., Maga, G., Jonsson, Z. O., May, R. P., Glatter, O. and Hubscher, U. (1997) The solution structure of functionally active human proliferating cell nuclear antigen determined by small-angle neutron scattering. *J Mol Biol*, 275, 123-32

Sever-Chroneos, Z., Angus, S. P., Fribourg, A. F., Wan, H., Todorov, I., Knudsen, K. E. and Knudsen, E. S. (2001) Retinoblastoma tumor suppressor protein signals through inhibition of cyclin-dependent kinase 2 activity to disrupt PCNA function in S phase. *Mol Cell Biol*, 21, 4032-45

Shiomi, Y., Shinozaki, A., Nakada, D., Sugimoto, K., Usukura, J., Obuse, C. and Tsurimoto, T. (2002) Clamp and clamp loader structures of the human checkpoint protein complexes, Rad9-1-1 and Rad17-RFC. *Genes Cells*, 7, 861-8

Shivji, M. K., Podust, V. N., Hubscher, U. and Wood, R. D. (1995) Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry*, 34, 5011-7

Tsurimoto, T. (1999) PCNA binding proteins. *Front Biosci*, 4, D849-58

Tsurimoto, T. and Stillman, B. (1989) Purification of a cellular replication factor, RF-C, that is required for coordinated synthesis of leading and lagging strands during simian virus 40 DNA replication in vitro. *Mol Cell Biol*, 9, 609-19

Tsurimoto, T. and Stillman, B. (1990) Functions of replication factor C and proliferating-cell nuclear antigen: functional similarity of DNA polymerase accessory proteins from human cells and bacteriophage T4. *Proc Natl Acad Sci U S A*, 87, 1023-7

Tsurimoto, T. and Stillman, B. (1991) Replication factors required for SV40 DNA replication in vitro. I. DNA structure-specific recognition of a primer-template junction by eukaryotic DNA polymerases and their accessory proteins. *J Biol Chem*, 266, 1950-60

Uhlmann, F., Cai, J., Flores-Rozas, H., Dean, F. B., Finkelstein, J., O'Donnell, M. and Hurwitz, J. (1996) In vitro reconstitution of human replication factor C from its five subunits. *Proc Natl Acad Sci U S A*, 93, 6521-6

Uhlmann, F., Cai, J., Gibbs, E., O'Donnell, M. and Hurwitz, J. (1997a) Deletion analysis of the large subunit p140 in human replication factor C reveals regions required for complex formation and replication activities. *J Biol Chem*, 272, 10058-64

Uhlmann, F., Gibbs, E., Cai, J., O'Donnell, M. and Hurwitz, J. (1997b) Identification of regions within the four small subunits of human replication factor C required for complex formation and DNA replication. *J Biol Chem*, 272, 10065-7

van Brabant, A. J., Stan, R. and Ellis, N. A. (2000) DNA helicases, genomic instability, and human genetic disease. *Annu Rev Genomics Hum Genet*, 1, 409-59

Venclovas, C., Colvin, M. E. and Thelen, M. P. (2002) Molecular modeling-based analysis of interactions in the RFC-dependent clamp-loading process. *Protein Sci*, 11, 2403-16

Waga, S. and Stillman, B. (1994) Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro. *Nature*, 369, 207-12

Wang, J. C. (2002) Cellular roles of DNA topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol*, 3, 430-40

Wold, M. S. (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem*, 66, 61-92

Wood, R. D. and Shivji, M. K. (1997) Which DNA polymerases are used for DNA-repair in eukaryotes? *Carcinogenesis*, 18, 605-10

Wyman, C. and Botchan, M. (1995) DNA replication. A familiar ring to DNA polymerase processivity. *Curr Biol*, 5, 334-7

Yoder, B. L. and Burgers, P. M. (1991) *Saccharomyces cerevisiae* replication factor C. I. Purification and characterization of its ATPase activity. *J Biol Chem*, 266, 22689-97

Zhang, G., Gibbs, E., Kelman, Z., O'Donnell, M. and Hurwitz, J. (1999) Studies on the interactions between human replication factor C and human proliferating cell nuclear antigen. *Proc Natl Acad Sci U S A*, 96, 1869-74

8 Abbreviations

Amp	Ampicillin
Ase I	Restriction enzyme from the bacterium *Aquaspirillum serpens*
BamH I	Restriction enzyme from the bacterium *Bacillus amyloliquefaciens H*
Bis-Tris	Bis-(2-hydroxyethyl)-amino-tris-(hydroxymethyl)-methane
bp	Base pair(s)
BSA	Bovine serum albumin
°C	Degrees centigrade
Cam	Chloramphenicol
cDNA	Complementary DNA
μCi	Micro-Curie (unit of measurement of radiactivity)
cpm	Counts per minute (radioactivity)
(d)ATP	(Deoxy)adenosine-5'-triphosphate
(d)CTP	(Deoxy)cytidine-5'-triphosphate
(d)GTP	(Deoxy)guanosine-5'-triphosphate
DNA	Deoxyribonucleic acid
dNTP	N = A (adenosine) + G (guanosine) + C (cytidine) + T (thymidine)
DTT	Dithiothreitol
(d)TTP	(Deoxy)thymidine-5'-triphosphate
Eag I	Restriction enzyme from the bacterium *Enterobacter agglomerans*
E. coli	*Escherichia coli*
EDTA	Ethylendiamintetraacetic acid
g	Gravity force (9.81 m/s^2)
[^{3}H]	Labelled with tritium
His	Histidine
Kan	Kanamycin
kDa	Kilodalton (= 1000 Dalton)
LB medium	Luria Bertani medium
M13	Bacteriophage which infects the bacterium *E. coli*
mRNA	Messenger RNA
Nde I	Restriction enzyme from the bacterium *Neisseria denitrificans*
nickel-IDA	Nickel-iminodiacetic acid (Ni^{2+} ions are chelated by three positions)
nickel-NTA	Nickel-nitrilotriacetic acid (Ni^{2+} ions are chelated by four positions)
NP-40	Nonidet P-40
OD_{600}	Optical density measured at a wavelength of 600 nm
[γ^{32}P]	Labelled with the γ-radiation-emitting ^{32}phosphorus isotope

PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
Pfu-polymerase	DNA polymerase from the bacterium *Pyrococcus furiosus*
PMSF	Phenylmethylsulfonyl fluoride
Pol	DNA polymerase
Rad	Radiation resistance (proteins involved in DNA checkpoint or repair)
RF-C	Replication factor C
RNA	Ribonucleic acid
RP-A	Replication protein A
Sal I	Restriction enzyme from the bacterium *Streptomyces albus G*
S. cerevisiae	*Saccharomyces cerevisiae* (budding yeast)
SDS	Sodiumdodecyl sulphate
SDS-PAGE	Sodiumdodecyl sulphate-polyacrylamide gel electrophoresis
Spe I	Restriction enzyme from the bacterium *Sphaerotilus species*
SSB	Single-strand DNA binding protein
TAE	Tris-acetate EDTA (40 mM Tris-acetate (pH 8.0), 1 mM EDTA)
TBE	Tris-borate EDTA (45 mM Tris-borate (pH 8.0), 1 mM EDTA)
TBS	Tris-buffered saline (10 mM Tris-HCl (pH 7.5), 150 mM NaCl)
TCA	Trichloroacetic acid
Tris	Tris-(hydroxymethyl)-aminomethane
tRNA	Transfer RNA
u	Units of enzyme activity
v/v	Volume per volume
w/v	Weight per volume
Xba I	Restriction enzyme from the bacterium *Xanthomonas badrii*
Xho I	Restriction enzyme from the bacterium *Xanthomonas holcicola*

Units of measurement:

		of amount of substance		*of concentration*		*of weight*		*of volume*	
1		**mole** ($6.02 \cdot 10^{23}$)	M	**molar**	g	**gramme**	L	**litre**	
10^{-3}	milli-	mmol	mM		mg		ml		
10^{-6}	mikro-	μmol	μM		μg		μl		
10^{-9}	nano-	nmol	nM		ng		nl		
10^{-12}	pico-	pmol	pM		pg		pl		

9 Acknowledgements

I would like to thank the following persons for their assistance and support:

Prof. Dr. Ulrich Hübscher for refereeing my thesis. I also wish to thank for his guidance and support and for the proofreading of my thesis work.

Prof. Dr. Hans Lutz for accepting to referee my thesis.

Erica Friedrich-Heineken for introducing me to cloning, expression and protein purification techniques. I would also like to thank for her excellent advice and the numerous constructive discussions. I especially appreciated her encouraging attitude.

Elena Ferrari for introducing me to the laboratory work. I very much appreciated her great experience with enzymatic assays as well as with the “tricky enzyme” RF-C.

Dr. Giovanni Maga for his scientific support and valuable advice thanks to his great knowledge of RF-C.

Dr. Kirsten Dahm for providing me the plasmids containing the genetic information for the RF-C subunits.

and all the other members of the Institute.

Special thanks go to my parents, my brother and Fabian Kälin for their moral support and interest in my work.

I would also like to thank the “Kantonale Drucksachen- und Materialzentrale Zürich” for supporting the printing of this thesis.

Finally, I would like to thank the University of Zürich for paying me a part time salary.

Curriculum vitae

Sibylle Kuonen

geboren am 7. August 1978 in Zürich
Heimatorte: Guttet-Feschel (VS) und Zürich

1985 – 1991	Primarschule in Zürich
1991 – 1998	Gymnasium an der Kantonsschule Freudenberg, Zürich, Matura Typus B
seit 1995	Zahlreiche kürzere und längere Aufenthalte in England und Wales, darunter ein einmonatiges Praktikum in einer Gross- und Kleintierklinik in Bodmin, Cornwall (2001)
1998 – 2003	Studium der Veterinärmedizin an der Universität Zürich
2001	Dreiwöchiges Praktikum am Institut für Veterinärbiochemie und Molekularbiologie
2001 – 2003	Dissertation am Institut für Veterinärbiochemie und Molekularbiologie der Universität Zürich unter der Leitung von Prof. Dr. U. Hübscher
2003	Eidgenössisches Staatsexamen der Veterinärmedizin in Zürich

November 2003